WEBVTT

00:00:00.000 --> 00:00:02.420 Support for Yale Cancer Answers

NOTE Confidence: 0.938980102539063

00:00:02.420 --> 00:00:04.840 comes from AstraZeneca, working

NOTE Confidence: 0.938980102539063

00:00:04.919 --> 00:00:07.259 side by side with leading

NOTE Confidence: 0.938980102539063

00:00:07.259 --> 00:00:09.599 scientists to better understand how

NOTE Confidence: 0.938980102539063

00:00:09.683 --> 00:00:15.611 complex data can be converted into

NOTE Confidence: 0.938980102539063

00:00:15.611 --> 00:00:16.599 innovative treatments. More information at astrazeneca-us.com.

NOTE Confidence: 0.938980102539063

00:00:16.600 --> 00:00:18.412 Welcome to Yale Cancer Answers with

NOTE Confidence: 0.938980102539063

00:00:18.412 --> 00:00:20.749 your host doctor Anees Chagpar.

NOTE Confidence: 0.938980102539063

00:00:20.749 --> 00:00:22.534 Yale Cancer Answers features the

NOTE Confidence: 0.938980102539063

00:00:22.534 --> 00:00:24.345 latest information on cancer care

NOTE Confidence: 0.938980102539063

00:00:24.345 --> 00:00:25.729 by welcoming oncologists and

NOTE Confidence: 0.938980102539063

00:00:25.729 --> 00:00:27.436 specialists who are on the

NOTE Confidence: 0.938980102539063

00:00:27.436 --> 00:00:29.434 forefront of the battle to fight

NOTE Confidence: 0.938980102539063

00:00:29.434 --> 00:00:31.190 cancer. This week it's a

NOTE Confidence: 0.938980102539063

00:00:31.190 --> 00:00:32.570 conversation about deep learning

NOTE Confidence: 0.938980102539063

00:00:32.570 --> 00:00:34.319 and cancer outcomes with Doctor

NOTE Confidence: 0.938980102539063

00:00:34.319 --> 00:00:36.702 Sanjay Aneja. Doctor Aneja is

NOTE Confidence: 0.938980102539063

00:00:36.702 --> 00:00:38.154 an assistant professor of

NOTE Confidence: 0.938980102539063

00:00:38.154 --> 00:00:39.802 therapeutic radiology at the Yale

NOTE Confidence: 0.938980102539063

00:00:39.802 --> 00:00:41.538 School of Medicine, where doctor

NOTE Confidence: 0.938980102539063

00:00:41.538 --> 00:00:44.142 Chagpar is a professor of

NOTE Confidence: 0.938980102539063

00:00:44.142 --> 00:00:45.310 surgical oncology.

NOTE Confidence: 0.917408168315887

00:00:45.310 --> 00:00:46.815 Maybe we can start

NOTE Confidence: 0.917408168315887

00:00:46.815 --> 00:00:48.978 off by you telling us a little

NOTE Confidence: 0.917408168315887

00:00:48.978 --> 00:00:50.478 bit about yourself and about

NOTE Confidence: 0.917408168315887

00:00:50.480 --> 00:00:51.380 your research?

NOTE Confidence: 0.917408168315887

00:00:51.380 --> 00:00:53.830 Sure, I've been in New Haven since 2009.

NOTE Confidence: 0.917408168315887

00:00:53.830 --> 00:00:55.612 I was actually a Yale medical

NOTE Confidence: 0.917408168315887

00:00:55.612 --> 00:00:57.470 student and I stayed here for

NOTE Confidence: 0.917408168315887

00:00:57.470 --> 00:00:59.294 my residency and now on faculty.

NOTE Confidence: 0.917408168315887

00:00:59.300 --> 00:01:00.795 Clinically, I am a radiation

NOTE Confidence: 0.917408168315887

00:01:00.795 --> 00:01:02.290 Oncologist and I primarily treat

NOTE Confidence: 0.917408168315887

00:01:02.344 --> 00:01:04.156 tumors of the central nervous system.

NOTE Confidence: 0.917408168315887

00:01:04.160 --> 00:01:05.680 So brain tumors as well

NOTE Confidence: 0.917408168315887

00:01:05.680 --> 00:01:06.896 as some prostate cancer.

NOTE Confidence: 0.917408168315887

00:01:06.900 --> 00:01:09.042 But I also run a large research group which

NOTE Confidence: 0.917408168315887

00:01:09.042 --> 00:01:11.460 is primarily focused on Applied Mathematics.

NOTE Confidence: 0.917408168315887

00:01:11.460 --> 00:01:13.320 My background is in applied mathematics, it's

NOTE Confidence: 0.917408168315887

00:01:13.320 --> 00:01:15.469 always something that I was very interested in.

NOTE Confidence: 0.917408168315887

00:01:15.470 --> 00:01:17.295 And what we're particularly interested

NOTE Confidence: 0.917408168315887

00:01:17.295 --> 00:01:20.401 in in my lab is looking at the utility

NOTE Confidence: 0.917408168315887

00:01:20.401 --> 00:01:22.025 of machine learning techniques,

NOTE Confidence: 0.917408168315887

00:01:22.030 --> 00:01:23.510 specifically deep learning in

NOTE Confidence: 0.917408168315887

00:01:23.510 --> 00:01:24.990 improving cancer outcomes and

NOTE Confidence: 0.917408168315887

00:01:24.990 --> 00:01:26.280 modeling cancer processes.

NOTE Confidence: 0.917408168315887

00:01:26.280 --> 00:01:27.820 I was with you

NOTE Confidence: 0.932632088661194

00:01:27.820 --> 00:01:30.136 all the way

NOTE Confidence: 0.932632088661194

00:01:30.136 --> 00:01:31.680 up to applied mathematics

NOTE Confidence: 0.932632088661194

00:01:31.680 --> 00:01:33.870 and machine learning and deep learning

NOTE Confidence: 0.932632088661194

00:01:33.870 --> 00:01:37.088 and all of that sounds really deep.

NOTE Confidence: 0.932632088661194

00:01:37.090 --> 00:01:39.430 But can you break it down

NOTE Confidence: 0.932632088661194

00:01:39.430 --> 00:01:41.330 into simple terms for us?

NOTE Confidence: 0.932632088661194

00:01:41.330 --> 00:01:44.030 What exactly are you doing?

NOTE Confidence: 0.935569607294523

00:01:44.030 --> 00:01:45.346 That's a good question,

NOTE Confidence: 0.935569607294523

00:01:45.346 --> 00:01:47.953 I think that we're very interested in that

NOTE Confidence: 0.935569607294523

00:01:47.953 --> 00:01:50.249 we have such a large amount of healthcare

NOTE Confidence: 0.935569607294523

00:01:50.317 --> 00:01:52.549 data that's currently available to us

NOTE Confidence: 0.935569607294523

00:01:52.550 --> 00:01:54.296 that's been kind of cultivated over

NOTE Confidence: 0.935569607294523

00:01:54.296 --> 00:01:55.950 years of managing cancer patients,

NOTE Confidence: 0.935569607294523

00:01:55.950 --> 00:01:58.134 and what we'd like to do is develop

NOTE Confidence: 0.935569607294523

00:01:58.134 --> 00:02:00.078 methods to better model that data so

NOTE Confidence: 0.935569607294523

00:02:00.078 --> 00:02:02.598 that we can kind of use that information

NOTE Confidence: 0.935569607294523

00:02:02.598 --> 00:02:04.842 to better improve the health care

NOTE Confidence: 0.935569607294523

00:02:04.842 --> 00:02:06.760 outcomes for cancer patients moving forward.

NOTE Confidence: 0.935569607294523

00:02:06.760 --> 00:02:08.265 Machine learning is just one

NOTE Confidence: 0.935569607294523

00:02:08.265 --> 00:02:10.160 way in which we do that.

NOTE Confidence: 0.935569607294523

00:02:10.160 --> 00:02:10.480 Traditionally,

NOTE Confidence: 0.935569607294523

00:02:10.480 --> 00:02:12.720 the way in which we used to

NOTE Confidence: 0.935569607294523

00:02:12.720 --> 00:02:14.179 model cancer as a disease

NOTE Confidence: 0.935569607294523

00:02:14.180 --> 00:02:16.364 it's always been a very difficult process.

NOTE Confidence: 0.935569607294523

00:02:16.370 --> 00:02:18.656 We would look at factors

NOTE Confidence: 0.935569607294523

00:02:18.656 --> 00:02:20.889 that physicians thought were important,

NOTE Confidence: 0.935569607294523

00:02:20.890 --> 00:02:22.986 and then we put them in a model

NOTE Confidence: 0.935569607294523

00:02:22.986 --> 00:02:25.349 and then we kind of look at an

NOTE Confidence: 0.935569607294523

00:02:25.349 --> 00:02:27.849 average and that had a couple

NOTE Confidence: 0.935569607294523

00:02:27.849 --> 00:02:29.939 different problems associated with it.

NOTE Confidence: 0.935569607294523

00:02:29.940 --> 00:02:31.716 One is that physicians aren't really

NOTE Confidence: 0.935569607294523

00:02:31.716 --> 00:02:33.315 great at predicting what factors

NOTE Confidence: 0.935569607294523

00:02:33.315 --> 00:02:35.100 are actually associated with cancer.

NOTE Confidence: 0.935569607294523

00:02:35.100 --> 00:02:36.408 It's a really complex disease.

NOTE Confidence: 0.935569607294523

00:02:36.408 --> 00:02:38.370 The second thing is that cancer

NOTE Confidence: 0.935569607294523

00:02:38.427 --> 00:02:39.947 is very difficult to model,

NOTE Confidence: 0.935569607294523

00:02:39.950 --> 00:02:42.310 and so using some of these techniques that

NOTE Confidence: 0.935569607294523

00:02:42.310 --> 00:02:44.790 we were developing a number of years ago,

NOTE Confidence: 0.935569607294523

00:02:44.790 --> 00:02:46.134 those techniques weren't

NOTE Confidence: 0.935569607294523

00:02:46.134 --> 00:02:47.478 necessarily as effective.

NOTE Confidence: 0.935569607294523

00:02:47.480 --> 00:02:49.106 Machine learning is sort of an

NOTE Confidence: 0.935569607294523

00:02:49.106 --> 00:02:50.480 advanced form of modeling data.

NOTE Confidence: 0.935569607294523

00:02:50.480 --> 00:02:52.568 What it does is it takes in all

NOTE Confidence: 0.935569607294523

00:02:52.568 --> 00:02:53.490 types of data,

NOTE Confidence: 0.935569607294523

00:02:53.490 --> 00:02:55.158 so it doesn't really require the

NOTE Confidence: 0.935569607294523

00:02:55.158 --> 00:02:57.109 physician to make any sort of choices

NOTE Confidence: 0.935569607294523

00:02:57.109 --> 00:02:58.943 about what type of data to include,

NOTE Confidence: 0.935569607294523

00:02:58.950 --> 00:03:00.770 and it allows us to model very

NOTE Confidence: 0.935569607294523

00:03:00.770 --> 00:03:01.950 complex processes like cancer,

NOTE Confidence: 0.935569607294523

00:03:01.950 --> 00:03:03.693 and there's been a lot of effort

NOTE Confidence: 0.935569607294523

00:03:03.693 --> 00:03:05.673 within our group to show that machine

NOTE Confidence: 0.935569607294523

00:03:05.673 --> 00:03:07.437 learning methods are probably the best

NOTE Confidence: 0.935569607294523

00:03:07.493 --> 00:03:09.589 way in which we can model cancer outcomes,

NOTE Confidence: 0.935569607294523

00:03:09.590 --> 00:03:11.228 and so that's what we're particularly

NOTE Confidence: 0.929244518280029

00:03:11.230 --> 00:03:13.494 interested in.

NOTE Confidence: 0.929244518280029

00:03:13.494 --> 00:03:15.865 Can you give us an example of how you did that in your lab?

00:03:18.400 --> 00:03:20.620 The concept of taking large amounts of data variables

NOTE Confidence: 0.929244518280029

00:03:20.620 --> 00:03:22.810 of various sorts that physicians may

NOTE Confidence: 0.929244518280029

00:03:22.810 --> 00:03:25.378 or may not think are relevant at all,

NOTE Confidence: 0.929244518280029

00:03:25.380 --> 00:03:27.816 giving it to a computer and saying,

NOTE Confidence: 0.929244518280029

00:03:27.820 --> 00:03:30.916 hey, look at all of this data and see

NOTE Confidence: 0.929244518280029

00:03:30.916 --> 00:03:33.406 whether or not any of these things,

NOTE Confidence: 0.929244518280029

00:03:33.410 --> 00:03:35.559 or a combination of these things may

NOTE Confidence: 0.929244518280029

00:03:35.559 --> 00:03:37.599 actually predict a particular outcome.

NOTE Confidence: 0.929244518280029

00:03:37.600 --> 00:03:40.036 Am I on the right track?

NOTE Confidence: 0.929244518280029

00:03:40.040 --> 00:03:40.740 Yeah, definitely.

00:03:41.090 --> 00:03:43.526 I think the one benefit of machine learning,

NOTE Confidence: 0.939530432224274

00:03:43.530 --> 00:03:45.530 compared to more traditional techniques

NOTE Confidence: 0.939530432224274

00:03:45.530 --> 00:03:48.392 of modeling cancer data is that it allows

NOTE Confidence: 0.939530432224274

00:03:48.392 --> 00:03:50.930 us to look at all the various resources.

NOTE Confidence: 0.939530432224274

00:03:52.640 --> 00:03:54.390 An example of a project that

NOTE Confidence: 0.939530432224274

00:03:54.390 --> 00:03:56.647 we've done in our lab is trying to

NOTE Confidence: 0.939530432224274

00:03:56.647 --> 00:03:58.092 model the outcomes for patients

NOTE Confidence: 0.939530432224274

00:03:58.155 --> 00:03:59.770 with early stage lung cancer.

NOTE Confidence: 0.939530432224274

00:03:59.770 --> 00:04:01.674 And so what we do in clinical practice

NOTE Confidence: 0.939530432224274

00:04:01.674 --> 00:04:03.655 is that we look at various different

NOTE Confidence: 0.939530432224274

00:04:03.655 --> 00:04:05.495 things when we're trying to model

NOTE Confidence: 0.939530432224274

00:04:05.495 --> 00:04:07.175 early stage lung cancer patients.

NOTE Confidence: 0.939530432224274

00:04:07.180 --> 00:04:08.610 We look at certain demographic

NOTE Confidence: 0.939530432224274

00:04:08.610 --> 00:04:10.644 variables like your age and if they're

NOTE Confidence: 0.939530432224274

00:04:10.644 --> 00:04:12.306 smoking and things of that nature,

NOTE Confidence: 0.939530432224274

00:04:12.310 --> 00:04:14.174 we also look at the images to see

NOTE Confidence: 0.939530432224274

00:04:14.174 --> 00:04:16.069 how big the tumor looks and whether

NOTE Confidence: 0.939530432224274

00:04:16.069 --> 00:04:18.237 or not it's close to any structures

NOTE Confidence: 0.939530432224274

00:04:18.237 --> 00:04:19.717 that we're worried about.

NOTE Confidence: 0.939530432224274

00:04:19.720 --> 00:04:21.496 And then we also look at

NOTE Confidence: 0.939530432224274

00:04:21.500 --> 00:04:23.236 what are treatment plans are

NOTE Confidence: 0.939530432224274

00:04:23.236 --> 00:04:25.350 and so how well we can deliver radiation

NOTE Confidence: 0.939530432224274

00:04:25.350 --> 00:04:27.577 to treat those and those are kind

NOTE Confidence: 0.939530432224274

00:04:27.577 --> 00:04:29.281 of three different data sources of

NOTE Confidence: 0.939530432224274

00:04:29.281 --> 00:04:31.509 sorts that we use in clinic in order

NOTE Confidence: 0.939530432224274

00:04:31.509 --> 00:04:33.327 to determine whether or not a patient

NOTE Confidence: 0.939530432224274

00:04:33.327 --> 00:04:35.220 will have a good outcome or bad outcome.

NOTE Confidence: 0.939530432224274

00:04:35.220 --> 00:04:36.648 And currently the only models that

NOTE Confidence: 0.939530432224274

00:04:36.648 --> 00:04:38.144 we actually have to tell patients

NOTE Confidence: 0.939530432224274

00:04:38.144 --> 00:04:39.626 how their outcomes would be are

NOTE Confidence: 0.939530432224274

00:04:39.626 --> 00:04:41.140 using only demographic variables,

NOTE Confidence: 0.939530432224274

00:04:41.140 --> 00:04:43.284 and so they're not really using the pictures,

NOTE Confidence: 0.939530432224274

00:04:43.290 --> 00:04:44.898 and they're not really using the

NOTE Confidence: 0.939530432224274

00:04:44.898 --> 00:04:45.702 treatment planning information,

NOTE Confidence: 0.939530432224274

00:04:45.710 --> 00:04:47.334 and the reason for that is because

NOTE Confidence: 0.939530432224274

00:04:47.334 --> 00:04:48.345 that data isn't necessarily

NOTE Confidence: 0.939530432224274

00:04:48.345 --> 00:04:49.929 something that you can put into

NOTE Confidence: 0.939530432224274

00:04:49.929 --> 00:04:51.659 some of those traditional models.

NOTE Confidence: 0.939530432224274

00:04:51.660 --> 00:04:53.740 What we did is we developed a deep

NOTE Confidence: 0.939530432224274

00:04:53.792 --> 00:04:55.620 learning machine learning algorithm.

NOTE Confidence: 0.939530432224274

00:04:55.620 --> 00:04:57.108 So it's an algorithm that takes

NOTE Confidence: 0.939530432224274

00:04:57.108 --> 00:04:59.129 the rawest form of the data from

NOTE Confidence: 0.939530432224274

00:04:59.129 --> 00:05:00.429 the electronic medical record,

NOTE Confidence: 0.939530432224274

00:05:00.430 --> 00:05:02.128 it pulls that demographic data from

NOTE Confidence: 0.939530432224274

00:05:02.128 --> 00:05:03.260 the electronic medical record,

NOTE Confidence: 0.939530432224274

00:05:03.260 --> 00:05:04.952 it takes every pixel from every

NOTE Confidence: 0.939530432224274

00:05:04.952 --> 00:05:06.694 picture of the tumor and analyzes

NOTE Confidence: 0.939530432224274

00:05:06.694 --> 00:05:08.633 those pixels in a very unique way.

NOTE Confidence: 0.939530432224274

00:05:08.640 --> 00:05:10.616 And then it also looks at every little

NOTE Confidence: 0.939530432224274

00:05:10.616 --> 00:05:12.117 part of our radiation treatment

NOTE Confidence: 0.939530432224274

00:05:12.117 --> 00:05:14.013 plan down to the pixel level.

NOTE Confidence: 0.939530432224274

00:05:14.020 --> 00:05:15.651 It kind of coalesces all that information

NOTE Confidence: 0.939530432224274

00:05:15.651 --> 00:05:17.409 and derives a personalized prediction,

NOTE Confidence: 0.939530432224274

00:05:17.410 --> 00:05:19.108 which we found was better than

NOTE Confidence: 0.939530432224274

00:05:19.594 --> 00:05:21.046 sort of getting an average based

NOTE Confidence: 0.939530432224274

00:05:21.046 --> 00:05:22.788 on just the demographic variables

NOTE Confidence: 0.915503346920013

00:05:22.790 --> 00:05:24.630 alone. So basically it's

NOTE Confidence: 0.915503346920013

00:05:24.630 --> 00:05:26.470 taking all of this information,

NOTE Confidence: 0.915503346920013

00:05:26.470 --> 00:05:28.455 the clinical information that

NOTE Confidence: 0.915503346920013

00:05:28.455 --> 00:05:30.440 most clinicians would use,

NOTE Confidence: 0.915503346920013

00:05:30.511 --> 00:05:32.499 the image Ng that they also use,

NOTE Confidence: 0.915503346920013

00:05:32.500 --> 00:05:34.786 but that they can't really put

NOTE Confidence: 0.915503346920013

00:05:34.786 --> 00:05:37.349 into a model because it's hard to

NOTE Confidence: 0.915503346920013

00:05:37.349 --> 00:05:39.596 define, like I see a big tumor,

NOTE Confidence: 0.915503346920013

00:05:39.600 --> 00:05:41.730 it looks like it's encasing

NOTE Confidence: 0.915503346920013

00:05:41.730 --> 00:05:42.795 some important vessels,

NOTE Confidence: 0.915503346920013

00:05:42.800 --> 00:05:45.208 but how do I really put that into

NOTE Confidence: 0.915503346920013

00:05:45.208 --> 00:05:47.353 a model and the treatment

NOTE Confidence: 0.915503346920013

00:05:47.353 --> 00:05:50.012 plan and can tell you kind of

NOTE Confidence: 0.915503346920013

00:05:50.012 --> 00:05:52.298 trying to be a clinician because

NOTE Confidence: 0.915503346920013

00:05:52.298 --> 00:05:54.918 a clinician will kind of look at

NOTE Confidence: 0.915503346920013

00:05:54.918 --> 00:05:57.829 that and have a good shift alt of.

NOTE Confidence: 0.915503346920013

00:05:57.830 --> 00:05:59.250 This patient will do well.

NOTE Confidence: 0.915503346920013

00:05:59.250 --> 00:06:01.238 This patient won't do so well and

NOTE Confidence: 0.915503346920013

00:06:01.238 --> 00:06:02.980 the computer can kind of give

NOTE Confidence: 0.915503346920013

00:06:02.980 --> 00:06:04.642 you that in a more quantitative

NOTE Confidence: 0.906355142593384

00:06:04.650 --> 00:06:05.786 way.

NOTE Confidence: 0.906355142593384

00:06:05.786 --> 00:06:07.290 Yeah, I think that you're touching

NOTE Confidence: 0.906355142593384

00:06:07.290 --> 00:06:08.910 upon one of the big advantages

NOTE Confidence: 0.906355142593384

00:06:08.910 --> 00:06:10.610 of these sorts of techniques.

NOTE Confidence: 0.906355142593384

00:06:10.610 --> 00:06:12.874 So one is that there an objective form,

NOTE Confidence: 0.906355142593384

00:06:12.880 --> 00:06:14.300 and so it's not necessarily

NOTE Confidence: 0.906355142593384

00:06:14.300 --> 00:06:15.436 utilizing one physicians experiences,

NOTE Confidence: 0.906355142593384

00:06:15.440 --> 00:06:16.684 or other physician experiences trying

NOTE Confidence: 0.906355142593384

00:06:16.684 --> 00:06:18.239 to use everyone's collective experience

NOTE Confidence: 0.906355142593384

00:06:18.239 --> 00:06:19.978 of analyzing data in an objective way.

NOTE Confidence: 0.906355142593384

00:06:19.980 --> 00:06:22.116 The other thing I think you're kind of

NOTE Confidence: 0.906355142593384

00:06:22.116 --> 00:06:23.888 touching upon is this idea of we're

NOTE Confidence: 0.906355142593384

00:06:23.888 --> 00:06:25.848 trying to mimic the same set of

NOTE Confidence: 0.906355142593384

00:06:25.848 --> 00:06:27.648 predictions that physicians make.

NOTE Confidence: 0.906355142593384

00:06:27.650 --> 00:06:29.372 And that's another reason why a lot

NOTE Confidence: 0.906355142593384

00:06:29.372 --> 00:06:31.168 of people are very interested in

NOTE Confidence: 0.906355142593384

00:06:31.168 --> 00:06:32.833 machine learning is because there

NOTE Confidence: 0.906355142593384

00:06:32.840 --> 00:06:34.404 is this component of artificial

NOTE Confidence: 0.906355142593384

00:06:34.404 --> 00:06:36.750 intelligence that can be kind of

NOTE Confidence: 0.906355142593384

00:06:36.814 --> 00:06:39.012 created when you are able to look

NOTE Confidence: 0.906355142593384

00:06:39.012 --> 00:06:41.315 at data sources without

NOTE Confidence: 0.906355142593384

00:06:41.315 --> 00:06:43.475 choosing which variables to evaluate,

NOTE Confidence: 0.924597144126892

00:06:43.480 --> 00:06:45.760 and so

NOTE Confidence: 0.924597144126892

00:06:45.760 --> 00:06:47.926 in this project where you

NOTE Confidence: 0.924597144126892

00:06:47.926 --> 00:06:49.831 were looking at outcomes of

NOTE Confidence: 0.924597144126892

00:06:49.831 --> 00:06:52.159 early lung cancer and giving a

NOTE Confidence: 0.924597144126892

00:06:52.159 --> 00:06:53.740 machine the demographic data,

NOTE Confidence: 0.924597144126892

00:06:53.740 --> 00:06:56.020 the smoking data, the imaging data,

NOTE Confidence: 0.924597144126892

00:06:56.020 --> 00:06:57.448 the treatment plan data,

NOTE Confidence: 0.924597144126892

00:06:57.448 --> 00:07:00.393 and you found that it was able to

NOTE Confidence: 0.924597144126892

00:07:00.393 --> 00:07:02.577 predict outcomes in terms of survival

NOTE Confidence: 0.924597144126892

00:07:02.577 --> 00:07:05.090 or in terms of recurrence.

NOTE Confidence: 0.920851290225983

00:07:05.090 --> 00:07:07.367 Yes, we found that it was able to predict

NOTE Confidence: 0.920851290225983

00:07:07.367 --> 00:07:09.000 recurrence in various different ways,

NOTE Confidence: 0.920851290225983

00:07:09.000 --> 00:07:10.524 survival, and when we compared it

NOTE Confidence: 0.920851290225983

00:07:10.524 --> 00:07:11.897 to maybe just using traditional

NOTE Confidence: 0.920851290225983

00:07:11.897 --> 00:07:13.739 methods or just one data stream,

NOTE Confidence: 0.920851290225983

00:07:13.740 --> 00:07:15.777 we found that it outperformed all of

NOTE Confidence: 0.920851290225983

00:07:15.777 --> 00:07:17.539 those different methods and so this

NOTE Confidence: 0.920851290225983

00:07:17.539 --> 00:07:18.924 idea of combining everything together

NOTE Confidence: 0.920851290225983

00:07:18.924 --> 00:07:20.988 is very very essential and we

NOTE Confidence: 0.920851290225983

00:07:20.990 --> 00:07:22.808 know it's very intuitive for

NOTE Confidence: 0.920851290225983

00:07:22.808 --> 00:07:24.897 clinicians to realize you have to do that,

NOTE Confidence: 0.920851290225983

00:07:24.900 --> 00:07:26.574 but I think that it's important

NOTE Confidence: 0.920851290225983

00:07:26.574 --> 00:07:27.690 to be able to

NOTE Confidence: 0.920851290225983

00:07:27.690 --> 00:07:29.914 do that in a mathematical way as well.

NOTE Confidence: 0.920851290225983

00:07:29.920 --> 00:07:31.032 Did it outperform the

NOTE Confidence: 0.920851290225983

00:07:31.032 --> 00:07:32.422 best guess of a clinician?

NOTE Confidence: 0.920851290225983

00:07:32.430 --> 00:07:34.098 So instead of looking at just

NOTE Confidence: 0.920851290225983

00:07:34.098 --> 00:07:35.500 traditional models, we know that

NOTE Confidence: 0.920851290225983

00:07:35.500 --> 00:07:36.900 clinicians sometimes bring their own

NOTE Confidence: 0.920851290225983

00:07:36.900 --> 00:07:38.646 experience and expertise to the equation.

NOTE Confidence: 0.920851290225983

00:07:38.650 --> 00:07:40.110 Did you compare the machine

NOTE Confidence: 0.920851290225983

00:07:40.110 --> 00:07:41.856 learning to clinicians best guess,

NOTE Confidence: 0.920851290225983

00:07:41.860 --> 00:07:44.196 how well patients would do or not do?

NOTE Confidence: 0.920851290225983

00:07:44.200 --> 00:07:46.237 Yeah, so one thing that we've done

NOTE Confidence: 0.915704250335693

00:07:46.240 --> 00:07:47.107 is we've

NOTE Confidence: 0.915704250335693

00:07:47.107 --> 00:07:48.552 done studies that are

NOTE Confidence: 0.915704250335693

00:07:48.552 --> 00:07:50.042 looking at that, and similarly,

NOTE Confidence: 0.915704250335693

00:07:50.042 --> 00:07:51.788 what we've done is had multiple

NOTE Confidence: 0.915704250335693

00:07:51.788 --> 00:07:53.249 physicians do their best guess,

NOTE Confidence: 0.915704250335693

00:07:53.250 --> 00:07:55.126 and the first thing that's important to

NOTE Confidence: 0.915704250335693

00:07:55.126 --> 00:07:57.626 know is that physicians don't guess the same,

NOTE Confidence: 0.915704250335693

00:07:57.630 --> 00:07:59.464 and so there's actually not a gold

NOTE Confidence: 0.915704250335693

00:07:59.464 --> 00:08:01.889 standard for a way in which a physician

NOTE Confidence: 0.915704250335693

00:08:01.889 --> 00:08:03.469 would actually evaluate a patient.

NOTE Confidence: 0.915704250335693

00:08:03.470 --> 00:08:05.486 And we found that it performed at least

NOTE Confidence: 0.915704250335693

00:08:05.486 --> 00:08:07.557 as well as an experienced clinician,

NOTE Confidence: 0.915704250335693

00:08:07.560 --> 00:08:08.574 and better than

NOTE Confidence: 0.915704250335693

00:08:08.574 --> 00:08:10.280 maybe less experienced clinicians.

00:08:12.050 --> 00:08:13.820 In a sense,

NOTE Confidence: 0.925095915794373

00:08:13.820 --> 00:08:15.530 you're recreating with this machine

NOTE Confidence: 0.925095915794373

00:08:15.530 --> 00:08:17.240 learning the predictive

NOTE Confidence: 0.925095915794373

00:08:17.292 --> 00:08:19.127 ability of an experienced clinician.

NOTE Confidence: 0.925095915794373

00:08:19.130 --> 00:08:21.608 So how is that now being utilized?

NOTE Confidence: 0.925095915794373

00:08:21.610 --> 00:08:24.434 Or is it being utilized in the clinic?

NOTE Confidence: 0.925095915794373

00:08:24.440 --> 00:08:26.666 I think that one of the

NOTE Confidence: 0.925095915794373

00:08:26.666 --> 00:08:28.689 benefits of the platform that

NOTE Confidence: 0.925095915794373

00:08:28.690 --> 00:08:30.970 we've developed is that it doesn't

NOTE Confidence: 0.925095915794373

00:08:30.970 --> 00:08:33.524 actually require us to pull data and

NOTE Confidence: 0.925095915794373

00:08:33.524 --> 00:08:35.764 put it into a calculator of sorts,

NOTE Confidence: 0.925095915794373

00:08:35.770 --> 00:08:37.926 which is a lot of what we

NOTE Confidence: 0.925095915794373

00:08:37.926 --> 00:08:40.478 see with a lot of predictive

NOTE Confidence: 0.925095915794373

00:08:40.480 --> 00:08:41.275 things in cancer.

NOTE Confidence: 0.925095915794373

00:08:41.275 --> 00:08:42.865 And so what we're interested in right

NOTE Confidence: 0.925095915794373

00:08:42.865 --> 00:08:44.857 now is trying to connect ours to

NOTE Confidence: 0.925095915794373

00:08:44.857 --> 00:08:45.973 the electronic medical record.

NOTE Confidence: 0.925095915794373

00:08:45.980 --> 00:08:47.435 We've developed an iPhone application

NOTE Confidence: 0.925095915794373

00:08:47.435 --> 00:08:49.656 which allows us to basically put

NOTE Confidence: 0.925095915794373

00:08:49.656 --> 00:08:51.760 in the medical record number of a patient,

NOTE Confidence: 0.925095915794373

00:08:51.760 --> 00:08:54.100 and then it allows us to pull the data

NOTE Confidence: 0.925095915794373

00:08:54.100 --> 00:08:56.165 natively and then it allows us to kind

NOTE Confidence: 0.925095915794373

00:08:56.165 --> 00:08:58.357 of develop that prediction in the clinic,

NOTE Confidence: 0.925095915794373

00:08:58.360 --> 00:08:59.750 and that's the next step

NOTE Confidence: 0.925095915794373

00:08:59.750 --> 00:09:01.380 of what we're trying to do.

NOTE Confidence: 0.925095915794373

00:09:01.380 --> 00:09:03.193 But I think the other thing that's

NOTE Confidence: 0.925095915794373

00:09:03.193 --> 00:09:04.317 really important whenever we're

NOTE Confidence: 0.925095915794373

00:09:04.317 --> 00:09:05.827 thinking about these machine learning

NOTE Confidence: 0.925095915794373

00:09:05.827 --> 00:09:07.574 algorithms is because they are so

NOTE Confidence: 0.925095915794373

00:09:07.574 --> 00:09:08.804 good at modeling healthcare data,

NOTE Confidence: 0.925095915794373

00:09:08.810 --> 00:09:10.520 they tend to actually model it

NOTE Confidence: 0.925095915794373

00:09:10.520 --> 00:09:13.040 too well, and so that's what we call overfit.

NOTE Confidence: 0.925095915794373

00:09:13.040 --> 00:09:13.769 The data set,

NOTE Confidence: 0.925095915794373

00:09:13.769 --> 00:09:15.859 and so they sort of are very good

NOTE Confidence: 0.925095915794373

00:09:15.859 --> 00:09:17.797 at modeling Yale data for example,

NOTE Confidence: 0.925095915794373

00:09:17.800 --> 00:09:19.704 but they might not be so good

NOTE Confidence: 0.925095915794373

00:09:19.704 --> 00:09:20.880 at modeling data from

NOTE Confidence: 0.925095915794373

00:09:20.880 --> 00:09:21.804 I don't know Chicago,

NOTE Confidence: 0.925095915794373

00:09:21.804 --> 00:09:23.527 and what we're trying to do

NOTE Confidence: 0.925095915794373

00:09:23.527 --> 00:09:24.962 also is something called external

NOTE Confidence: 0.925095915794373

00:09:24.962 --> 00:09:26.787 validation where we send our model

NOTE Confidence: 0.925095915794373

00:09:26.787 --> 00:09:28.031 to different cancer

NOTE Confidence: 0.925095915794373

00:09:28.031 --> 00:09:29.840 centers across the country and say,

00:09:30.120 --> 00:09:32.360 don't even tell us what the outcomes are,

NOTE Confidence: 0.925095915794373

00:09:32.360 --> 00:09:34.320 we will tell you what our models are predicting,

NOTE Confidence: 0.925095915794373

00:09:34.320 --> 00:09:36.301 and then you tell us how good

NOTE Confidence: 0.925095915794373

00:09:36.301 --> 00:09:37.679 our model did so far.

NOTE Confidence: 0.925095915794373

00:09:37.680 --> 00:09:38.824 Our collaboration with Jefferson

NOTE Confidence: 0.925095915794373

00:09:38.824 --> 00:09:40.254 in Philadelphia has shown that

NOTE Confidence: 0.925095915794373

00:09:40.254 --> 00:09:41.670 the model is very productive.

NOTE Confidence: 0.925095915794373

00:09:41.670 --> 00:09:43.038 And it's maintaining that same performance,

NOTE Confidence: 0.925095915794373

00:09:43.040 --> 00:09:44.965 but it's important to

NOTE Confidence: 0.925095915794373

00:09:44.965 --> 00:09:46.814 test these models before we actually

NOTE Confidence: 0.925095915794373

00:09:46.814 --> 00:09:48.434 put them into clinical practice.

NOTE Confidence: 0.912570714950562

00:09:48.780 --> 00:09:51.836 How well something can

NOTE Confidence: 0.912570714950562

00:09:51.836 --> 00:09:55.295 predict is based on how well we learned and on

NOTE Confidence: 0.912570714950562

00:09:55.295 --> 00:09:58.037 the learning set that it had to work with.

NOTE Confidence: 0.912570714950562

00:09:58.040 --> 00:10:00.196 It makes sense, however, that the data

NOTE Confidence: 0.912570714950562

00:10:00.196 --> 00:10:02.849 that it was getting was objective data.

NOTE Confidence: 0.912570714950562

00:10:02.850 --> 00:10:05.114 It wasn't getting

NOTE Confidence: 0.912570714950562

00:10:05.114 --> 00:10:07.330 data that may have had a lot of

NOTE Confidence: 0.912570714950562

00:10:07.399 --> 00:10:09.703 factors that were subjective, right?

NOTE Confidence: 0.912570714950562

00:10:09.703 --> 00:10:11.418 You were looking at imaging.

NOTE Confidence: 0.912570714950562

00:10:11.420 --> 00:10:14.607 Well, the image is what it is and if you give

NOTE Confidence: 0.912570714950562

00:10:14.607 --> 00:10:17.247 that image to two different radiologists,

NOTE Confidence: 0.912570714950562

00:10:17.250 --> 00:10:18.758 they both should say

NOTE Confidence: 0.912570714950562

00:10:18.758 --> 00:10:20.266 roughly the same thing,

NOTE Confidence: 0.912570714950562

00:10:20.270 --> 00:10:22.230 maybe not exactly, but roughly,

NOTE Confidence: 0.912570714950562

00:10:22.230 --> 00:10:25.542 and so you'd think that the

NOTE Confidence: 0.912570714950562

00:10:25.542 --> 00:10:27.698 Jefferson images are going to be

NOTE Confidence: 0.912570714950562

00:10:27.700 --> 00:10:30.652 very much like Yale images and

NOTE Confidence: 0.912570714950562

00:10:30.652 --> 00:10:33.590 so that may account for that close

NOTE Confidence: 0.912570714950562

00:10:33.590 --> 00:10:36.570 correlation between the two datasets.

NOTE Confidence: 0.912570714950562

00:10:36.570 --> 00:10:38.646 But the next question is OK,

NOTE Confidence: 0.912570714950562

00:10:38.650 --> 00:10:40.883 let's suppose that the model after you

NOTE Confidence: 0.912570714950562

00:10:40.883 --> 00:10:43.389 test it and I don't want to

NOTE Confidence: 0.912570714950562

00:10:43.389 --> 00:10:45.404 minimize the utility of making sure

NOTE Confidence: 0.912570714950562

00:10:45.404 --> 00:10:47.300 that it's externally generalizable.

NOTE Confidence: 0.912570714950562

00:10:47.300 --> 00:10:50.060 Is it even being used here at Yale,

NOTE Confidence: 0.912570714950562

00:10:50.060 --> 00:10:52.153 where it was developed and

NOTE Confidence: 0.912570714950562

00:10:52.153 --> 00:10:54.680 it does well in terms of predicting

NOTE Confidence: 0.912570714950562

00:10:54.680 --> 00:10:57.680 outcomes as well as an experienced clinician?

NOTE Confidence: 0.912570714950562

00:10:57.680 --> 00:11:00.095 Is that being used in the clinic?

NOTE Confidence: 0.912570714950562

00:11:00.100 --> 00:11:02.522 Are you putting in this data

NOTE Confidence: 0.912570714950562

00:11:02.522 --> 00:11:03.560 you've got now,

NOTE Confidence: 0.912570714950562

00:11:03.560 --> 00:11:05.260 this iPhone application that can

NOTE Confidence: 0.912570714950562

00:11:05.260 --> 00:11:07.709 pull in this data into this model?

NOTE Confidence: 0.912570714950562

00:11:07.710 --> 00:11:11.013 The model can do its magic and tell you,

NOTE Confidence: 0.912570714950562

00:11:11.020 --> 00:11:12.938 this is the recurrence rate.

NOTE Confidence: 0.912570714950562

00:11:12.940 --> 00:11:14.310 This is the survival rate.

NOTE Confidence: 0.912570714950562

00:11:14.310 --> 00:11:16.228 Are you using that in the clinic,

NOTE Confidence: 0.912570714950562

00:11:16.230 --> 00:11:17.326 and if so how?

00:11:18.140 --> 00:11:19.784 I think that we're in the process

NOTE Confidence: 0.93610030412674

00:11:19.784 --> 00:11:20.880 of developing the application.

NOTE Confidence: 0.93610030412674

00:11:20.880 --> 00:11:22.880 One of the big hurdles and with health

NOTE Confidence: 0.93610030412674

00:11:22.880 --> 00:11:24.854 care in general is the ability to

NOTE Confidence: 0.93610030412674

00:11:24.854 --> 00:11:26.601 actually get access to the electronic

NOTE Confidence: 0.93610030412674

00:11:26.601 --> 00:11:28.540 medical record in a way in which

NOTE Confidence: 0.93610030412674

00:11:28.540 --> 00:11:30.196 you can make an application kind

NOTE Confidence: 0.93610030412674

00:11:30.196 --> 00:11:31.566 of seamlessly integrate into it.

NOTE Confidence: 0.93610030412674

00:11:31.570 --> 00:11:32.940 And so it's somewhat difficult

NOTE Confidence: 0.93610030412674

00:11:32.940 --> 00:11:34.310 for us to do that.

NOTE Confidence: 0.93610030412674

00:11:34.310 --> 00:11:35.948 We're working with

NOTE Confidence: 0.93610030412674

00:11:35.950 --> 00:11:37.046 a software engineering firm

NOTE Confidence: 0.93610030412674

00:11:37.046 --> 00:11:38.416 to actually help us with

NOTE Confidence: 0.93610030412674

00:11:38.420 --> 00:11:40.191 beyond the scope

NOTE Confidence: 0.93610030412674

00:11:40.191 --> 00:11:42.269 of what our lab does typically.

NOTE Confidence: 0.93610030412674

00:11:42.270 --> 00:11:43.950 And so that's where we're

NOTE Confidence: 0.93610030412674

00:11:43.950 --> 00:11:46.084 at right now with respect to actually

NOTE Confidence: 0.93610030412674

00:11:46.084 --> 00:11:47.412 integrating into clinical practice.

NOTE Confidence: 0.93610030412674

00:11:47.420 --> 00:11:49.128 I think that right now what we

NOTE Confidence: 0.93610030412674

00:11:49.128 --> 00:11:51.086 have is we have an ability to

NOTE Confidence: 0.93610030412674

00:11:51.086 --> 00:11:52.832 kind of look back on patients,

NOTE Confidence: 0.93610030412674

00:11:52.840 --> 00:11:55.010 and if there was a patient, for example,

NOTE Confidence: 0.93610030412674

00:11:55.010 --> 00:11:56.630 who wanted to have a prediction,

NOTE Confidence: 0.93610030412674

00:11:56.630 --> 00:11:57.990 we could actually generate that.

NOTE Confidence: 0.93610030412674

00:11:57.990 --> 00:11:59.887 But we cannot do it in the

NOTE Confidence: 0.918058693408966

00:11:59.890 --> 00:12:00.700 electronic medical record.

NOTE Confidence: 0.918058693408966

00:12:00.700 --> 00:12:01.780 As of right now,

00:12:03.140 --> 00:12:04.742 and so when you think about

NOTE Confidence: 0.918058693408966

00:12:04.742 --> 00:12:06.120 the potential utility of this,

NOTE Confidence: 0.918058693408966

00:12:06.120 --> 00:12:07.470 where do you see it

NOTE Confidence: 0.918058693408966

00:12:07.470 --> 00:12:08.466 going?

NOTE Confidence: 0.918058693408966

00:12:08.466 --> 00:12:10.725 So I think that one thing that I think

NOTE Confidence: 0.918058693408966

00:12:10.725 --> 00:12:12.765 is very important is as we're kind of

NOTE Confidence: 0.918058693408966

00:12:12.765 --> 00:12:14.340 developing so many different genres

NOTE Confidence: 0.918058693408966

00:12:14.340 --> 00:12:15.920 of treatment for cancer patients,

NOTE Confidence: 0.918058693408966

00:12:15.920 --> 00:12:17.840 there's this increasing need for us

NOTE Confidence: 0.918058693408966

00:12:17.840 --> 00:12:19.710 to develop methods to risk stratify

NOTE Confidence: 0.918058693408966

00:12:19.710 --> 00:12:21.420 them and identify the highest risk.

NOTE Confidence: 0.918058693408966

00:12:21.420 --> 00:12:23.415 Patients who maybe would be benefiting

NOTE Confidence: 0.918058693408966

00:12:23.415 --> 00:12:24.780 from more aggressive treatment,

NOTE Confidence: 0.918058693408966

00:12:24.780 --> 00:12:25.698 more aggressive followup,

NOTE Confidence: 0.918058693408966

00:12:25.698 --> 00:12:26.310 and similarly,

NOTE Confidence: 0.918058693408966

00:12:26.310 --> 00:12:28.630 I think we've found with some types of

NOTE Confidence: 0.918058693408966

00:12:28.630 --> 00:12:30.725 cancers that maybe we've been a little

NOTE Confidence: 0.918058693408966

00:12:30.725 --> 00:12:32.565 bit too aggressive in our follow-up

NOTE Confidence: 0.918058693408966

00:12:32.565 --> 00:12:34.515 or too aggressive with our therapy

NOTE Confidence: 0.918058693408966

00:12:34.515 --> 00:12:36.372 and to risk stratify

NOTE Confidence: 0.918058693408966

00:12:36.372 --> 00:12:38.136 which of those patients would be

NOTE Confidence: 0.918058693408966

00:12:38.136 --> 00:12:39.858 most useful for certain intervention

NOTE Confidence: 0.918058693408966

00:12:39.858 --> 00:12:41.578 versus another one is something

NOTE Confidence: 0.918058693408966

00:12:41.578 --> 00:12:43.322 that I think are our algorithm

NOTE Confidence: 0.918058693408966

00:12:43.322 --> 00:12:44.980 or our platform is very useful for

NOTE Confidence: 0.918058693408966

00:12:44.980 --> 00:12:46.530 especifically for early stage

NOTE Confidence: 0.918058693408966

00:12:46.530 --> 00:12:47.346 lung cancer patients.

NOTE Confidence: 0.918058693408966

00:12:47.346 --> 00:12:48.706 There's currently a clinical trial

NOTE Confidence: 0.918058693408966

00:12:48.710 --> 00:12:50.438 evaluating whether or not those

NOTE Confidence: 0.918058693408966

00:12:50.438 --> 00:12:52.125 patients should get radiation and then

NOTE Confidence: 0.918058693408966

00:12:52.125 --> 00:12:53.595 additional treatment on top of that.

NOTE Confidence: 0.918058693408966

00:12:53.600 --> 00:12:54.960 Because there's this idea that

NOTE Confidence: 0.918058693408966

00:12:54.960 --> 00:12:55.776 potentially additional immunotherapy,

NOTE Confidence: 0.918058693408966

00:12:55.780 --> 00:12:57.140 for example, would be helpful

NOTE Confidence: 0.918058693408966

00:12:57.140 --> 00:12:58.500 for those patients and a large

NOTE Confidence: 0.918058693408966

00:12:58.500 --> 00:13:00.166 amount of them may not need that

NOTE Confidence: 0.918058693408966

00:13:00.166 --> 00:13:01.556 because they are already going

NOTE Confidence: 0.918058693408966

00:13:01.556 --> 00:13:03.116 to have great outcomes anyways,

NOTE Confidence: 0.918058693408966

00:13:03.120 --> 00:13:04.752 and a large amount of them

NOTE Confidence: 0.918058693408966

00:13:04.752 --> 00:13:05.840 maybe would need that,

NOTE Confidence: 0.918058693408966

00:13:05.840 --> 00:13:07.466 and they should get it

NOTE Confidence: 0.918058693408966

00:13:07.470 --> 00:13:08.830 maybe

NOTE Confidence: 0.918058693408966

00:13:08.830 --> 00:13:09.646 right after treatment,

NOTE Confidence: 0.918058693408966

00:13:09.650 --> 00:13:11.826 before we even know how the outcomes are,

NOTE Confidence: 0.918058693408966

00:13:11.830 --> 00:13:13.531 and so if we can identify which

NOTE Confidence: 0.918058693408966

00:13:13.531 --> 00:13:15.218 patients are those high risk patients

NOTE Confidence: 0.918058693408966

00:13:15.218 --> 00:13:16.733 versus those low risk patients,

NOTE Confidence: 0.918058693408966

00:13:16.740 --> 00:13:19.464 I think we could potentially tailor

NOTE Confidence: 0.918058693408966

00:13:19.464 --> 00:13:21.707 our treatments and better understand

NOTE Confidence: 0.918058693408966

00:13:21.707 --> 00:13:25.080 the way in which we can

NOTE Confidence: 0.914630115032196

00:13:25.080 --> 00:13:27.060 personalize care based on someone's

NOTE Confidence: 0.914630115032196

00:13:27.060 --> 00:13:28.644 images and everything else.

NOTE Confidence: 0.914630115032196

00:13:28.650 --> 00:13:31.026 I get that concept of,

NOTE Confidence: 0.914630115032196

00:13:31.030 --> 00:13:32.618 you know, risk stratification,

NOTE Confidence: 0.914630115032196

00:13:32.618 --> 00:13:34.206 especially for additional therapy.

NOTE Confidence: 0.914630115032196

00:13:34.210 --> 00:13:36.190 What would be interesting though,

NOTE Confidence: 0.914630115032196

00:13:36.190 --> 00:13:38.801 is to really look at how do

NOTE Confidence: 0.914630115032196

00:13:38.801 --> 00:13:40.960 patients do without any therapy?

NOTE Confidence: 0.914630115032196

00:13:40.960 --> 00:13:42.940 How do patients do with

NOTE Confidence: 0.914630115032196

00:13:42.940 --> 00:13:44.524 therapy X versus therapy Y?

NOTE Confidence: 0.914630115032196

00:13:44.530 --> 00:13:46.930 And how can we really

NOTE Confidence: 0.914630115032196

00:13:46.930 --> 00:13:47.730 personalize therapies

NOTE Confidence: 0.914630115032196

00:13:47.730 --> 00:13:50.450 given the data that we have?

NOTE Confidence: 0.914630115032196

00:13:50.450 --> 00:13:53.248 We are going to take a short break for a medical minute,

NOTE Confidence: 0.914630115032196

00:13:53.250 --> 00:13:55.452 but when we come back we will

NOTE Confidence: 0.914630115032196

00:13:55.452 --> 00:13:56.553 answer those questions.

NOTE Confidence: 0.914630115032196

00:13:56.560 --> 00:13:58.768 So stay tuned to learn more

NOTE Confidence: 0.914630115032196

00:13:58.768 --> 00:14:00.240 about deep learning in

NOTE Confidence: 0.9278564453125

00:14:00.240 --> 00:14:01.708 cancer outcomes with my

NOTE Confidence: 0.9278564453125

00:14:01.708 --> 00:14:03.543 guest doctor Sanjay Aneja.

NOTE Confidence: 0.9278564453125

00:14:03.550 --> 00:14:05.390 Support for Yale Cancer Answers

NOTE Confidence: 0.9278564453125

00:14:05.390 --> 00:14:06.494 comes from AstraZeneca,

NOTE Confidence: 0.9278564453125

00:14:06.500 --> 00:14:08.375 a biopharmaceutical business that

NOTE Confidence: 0.9278564453125

00:14:08.375 --> 00:14:10.777 is pushing the boundaries of science

NOTE Confidence: 0.9278564453125

00:14:10.777 --> 00:14:13.561 to deliver new cancer medicines. More

NOTE Confidence: 0.9278564453125

00:14:13.561 --> 00:14:14.953 information at astrazeneca-us.com.

NOTE Confidence: 0.9278564453125

00:14:14.960 --> 00:14:17.438 This is a medical minute about

NOTE Confidence: 0.9278564453125

00:14:17.438 --> 00:14:19.098 pancreatic cancer, which represents

NOTE Confidence: 0.9278564453125

00:14:19.098 --> 00:14:22.410 about 3% of all cancers in the US and

NOTE Confidence: 0.9278564453125

00:14:22.410 --> 00:14:24.480 about 7% of cancer deaths.

NOTE Confidence: 0.9278564453125

00:14:24.480 --> 00:14:26.655 Clinical trials are currently being

NOTE Confidence: 0.9278564453125

00:14:26.655 --> 00:14:28.395 offered at federally designated

NOTE Confidence: 0.9278564453125

00:14:28.395 --> 00:14:29.584 comprehensive cancer centers

NOTE Confidence: 0.9278564453125

00:14:29.584 --> 00:14:31.329 for the treatment of advanced

NOTE Confidence: 0.9278564453125

00:14:31.329 --> 00:14:32.929 stage and metastatic pancreatic

NOTE Confidence: 0.9278564453125

00:14:32.929 --> 00:14:34.420 cancer using chemotherapy

NOTE Confidence: 0.9278564453125

00:14:34.420 --> 00:14:36.490 and other novel therapies like FOLFIRINOX

NOTE Confidence: 0.9278564453125

00:14:36.490 --> 00:14:38.734 a combination of five

NOTE Confidence: 0.9278564453125

00:14:38.734 --> 00:14:40.230 different chemotherapies is the

NOTE Confidence: 0.9278564453125

00:14:40.295 --> 00:14:42.545 latest advances in the treatment

NOTE Confidence: 0.9278564453125

00:14:42.545 --> 00:14:44.345 of metastatic pancreatic cancer,

NOTE Confidence: 0.9278564453125

00:14:44.350 --> 00:14:45.520 and research continues

NOTE Confidence: 0.9278564453125

00:14:45.520 --> 00:14:47.470 in centers around the world

NOTE Confidence: 0.9278564453125

00:14:47.470 --> 00:14:49.470 looking into targeted therapies.

NOTE Confidence: 0.9278564453125

00:14:49.470 --> 00:14:51.550 And a recently discovered marker

NOTE Confidence: 0.9278564453125

00:14:51.550 --> 00:14:54.481 hENT 1. This has been a medical

NOTE Confidence: 0.9278564453125

00:14:54.481 --> 00:14:57.183 minute brought to you as a public

NOTE Confidence: 0.9278564453125

00:14:57.183 --> 00:14:59.309 service by Yale Cancer Center.

NOTE Confidence: 0.9278564453125

00:14:59.310 --> 00:15:01.860 More information is available at

NOTE Confidence: 0.9278564453125

00:15:01.860 --> 00:15:03.390 yalecancercenter.org. You're listening

NOTE Confidence: 0.9278564453125

00:15:03.390 --> 00:15:05.250 to Connecticut public radio.

NOTE Confidence: 0.9278564453125

00:15:05.250 --> 00:15:05.610 Welcome

NOTE Confidence: 0.926801145076752

00:15:05.610 --> 00:15:07.430 back to Yale Cancer Answers.

NOTE Confidence: 0.926801145076752

00:15:07.430 --> 00:15:09.670 This is doctor Anees Chagpar

NOTE Confidence: 0.926801145076752

00:15:09.670 --> 00:15:11.815 and I'm joined tonight by my

NOTE Confidence: 0.926801145076752

00:15:11.815 --> 00:15:14.047 guest doctor Sanjay Aneja we're

NOTE Confidence: 0.926801145076752

00:15:14.047 --> 00:15:16.032 discussing deep learning in cancer

NOTE Confidence: 0.926801145076752

00:15:16.032 --> 00:15:18.342 outcomes and right before the break,

NOTE Confidence: 0.926801145076752

00:15:18.350 --> 00:15:20.264 doctor Aneja was telling us

NOTE Confidence: 0.926801145076752

00:15:20.264 --> 00:15:22.870 about how he and his lab have

NOTE Confidence: 0.926801145076752

00:15:22.870 --> 00:15:24.538 really used machine learning.

NOTE Confidence: 0.926801145076752

00:15:24.540 --> 00:15:26.220 That is to say,

NOTE Confidence: 0.926801145076752

00:15:26.220 --> 00:15:27.900 Applied Mathematics and complicated

NOTE Confidence: 0.926801145076752

00:15:27.900 --> 00:15:29.827 computational models to really take in

NOTE Confidence: 0.926801145076752

00:15:29.827 --> 00:15:31.833 lots and lots of data that clinicians

NOTE Confidence: 0.926801145076752

00:15:31.833 --> 00:15:34.245 using their usual clinical guess to

NOTE Confidence: 0.926801145076752

00:15:34.245 --> 00:15:36.292 predict outcomes for cancer patients.

NOTE Confidence: 0.926801145076752

00:15:36.292 --> 00:15:39.204 And Sanjay, the example that

NOTE Confidence: 0.926801145076752

00:15:39.204 --> 00:15:41.779 you gave us in early lung cancer

NOTE Confidence: 0.926801145076752

00:15:41.780 --> 00:15:42.968 where you said,

NOTE Confidence: 0.926801145076752

00:15:42.968 --> 00:15:45.344 the demographics

NOTE Confidence: 0.926801145076752

00:15:45.350 --> 00:15:47.340 we took the imaging data,

NOTE Confidence: 0.926801145076752

00:15:47.340 --> 00:15:49.895 but then we took the treatment plan

NOTE Confidence: 0.926801145076752

00:15:49.895 --> 00:15:53.214 and we use that and we ask the computer

NOTE Confidence: 0.926801145076752

00:15:53.214 --> 00:15:56.532 to look at these things down to the

NOTE Confidence: 0.926801145076752

00:15:56.532 --> 00:15:59.244 pixel level and then predict outcomes.

NOTE Confidence: 0.926801145076752

00:15:59.250 --> 00:16:02.122 And it was very good at predicting those

NOTE Confidence: 0.926801145076752

00:16:02.122 --> 00:16:05.197 outcomes as good as an experienced clinician,

NOTE Confidence: 0.926801145076752

00:16:05.200 --> 00:16:06.940 and that's great, but

NOTE Confidence: 0.926801145076752

00:16:06.940 --> 00:16:08.906 my question is

NOTE Confidence: 0.926801145076752

00:16:08.906 --> 00:16:11.530 what happens if you

NOTE Confidence: 0.926801145076752

00:16:11.530 --> 00:16:12.484 don't?

NOTE Confidence: 0.926801145076752

00:16:12.484 --> 00:16:15.610 How do you get rid of the treatment

NOTE Confidence: 0.926801145076752

00:16:15.610 --> 00:16:18.410 part of that and predict outcomes

NOTE Confidence: 0.926801145076752

00:16:18.410 --> 00:16:20.810 without treatment to kind of get at

NOTE Confidence: 0.926801145076752

00:16:20.810 --> 00:16:23.400 the idea of, are we over treating

NOTE Confidence: 0.926801145076752

00:16:23.400 --> 00:16:24.140 some patients?

NOTE Confidence: 0.926801145076752

00:16:24.140 --> 00:16:26.737 Because if the computer doesn't have that,

NOTE Confidence: 0.926801145076752

00:16:26.740 --> 00:16:28.600 if all patients are treated,

NOTE Confidence: 0.926801145076752

00:16:28.600 --> 00:16:31.560 and that's the basis on which it learned,

NOTE Confidence: 0.926801145076752

00:16:31.560 --> 00:16:35.270 how do you take out one part of that model?

NOTE Confidence: 0.9190593957901

00:16:35.790 --> 00:16:37.240 That's a good question,

NOTE Confidence: 0.9190593957901

00:16:37.240 --> 00:16:38.690 and it's something that I

NOTE Confidence: 0.9190593957901

00:16:38.690 --> 00:16:40.140 think is not impossible.

NOTE Confidence: 0.9190593957901

00:16:40.140 --> 00:16:42.404 So I think that if we remove the

NOTE Confidence: 0.9190593957901

00:16:42.404 --> 00:16:44.487 treatment piece of it and just look at

NOTE Confidence: 0.9190593957901

00:16:44.487 --> 00:16:46.520 the images and the demographic data,

NOTE Confidence: 0.9190593957901

00:16:46.520 --> 00:16:47.692 basically pretreatment information,

NOTE Confidence: 0.9190593957901

00:16:47.692 --> 00:16:49.450 we find the model is actually

NOTE Confidence: 0.9190593957901

00:16:49.500 --> 00:16:50.580 quite predictive as well.

NOTE Confidence: 0.9190593957901

00:16:50.580 --> 00:16:52.284 It just improves significantly if we

NOTE Confidence: 0.9190593957901

00:16:52.284 --> 00:16:54.054 know exactly what types of treatments

NOTE Confidence: 0.9190593957901

00:16:54.054 --> 00:16:55.509 we provided for the patients,

NOTE Confidence: 0.9190593957901

00:16:55.510 --> 00:16:57.374 and so another example of a study that

NOTE Confidence: 0.9190593957901

00:16:57.374 --> 00:16:59.191 we've done which only uses pretreatment

NOTE Confidence: 0.9190593957901

00:16:59.191 --> 00:17:01.135 imaging has been evaluating lymph nodes

NOTE Confidence: 0.9190593957901

00:17:01.186 --> 00:17:02.756 in head and neck cancer patients.

NOTE Confidence: 0.9190593957901

00:17:02.760 --> 00:17:04.250 We were attempting to look

NOTE Confidence: 0.9190593957901

00:17:04.250 --> 00:17:06.040 at which lymph nodes we saw

NOTE Confidence: 0.9190593957901

00:17:06.040 --> 00:17:08.028 on CT imaging actually had the

NOTE Confidence: 0.9190593957901

00:17:08.028 --> 00:17:09.929 presence of cancer and we wanted to

NOTE Confidence: 0.9190593957901

00:17:09.929 --> 00:17:11.724 identify that so that maybe you know

NOTE Confidence: 0.9190593957901

00:17:11.724 --> 00:17:13.488 what we could do is more tailor

NOTE Confidence: 0.9190593957901

00:17:13.490 --> 00:17:15.226 the therapy for head and neck cancer

NOTE Confidence: 0.9190593957901

00:17:15.226 --> 00:17:16.679 patients. Oftentimes with head and neck cancer patients

NOTE Confidence: 0.9190593957901

00:17:16.680 --> 00:17:18.276 when we think that their lymph

NOTE Confidence: 0.9190593957901

00:17:18.276 --> 00:17:19.340 nodes don't have cancer,

NOTE Confidence: 0.9190593957901

00:17:19.340 --> 00:17:20.650 we have them undergo surgery

NOTE Confidence: 0.9190593957901

00:17:20.650 --> 00:17:22.270 and then we find

NOTE Confidence: 0.9190593957901

00:17:22.270 --> 00:17:23.600 those lymph nodes have cancer.

NOTE Confidence: 0.9190593957901

00:17:23.600 --> 00:17:24.925 They have to get radiation

NOTE Confidence: 0.9190593957901

00:17:24.925 --> 00:17:25.720 and chemotherapy altogether.

NOTE Confidence: 0.9190593957901

00:17:25.720 --> 00:17:27.640 And so if we were able to identify

NOTE Confidence: 0.9190593957901

00:17:27.640 --> 00:17:29.175 the patients ahead of time that

NOTE Confidence: 0.9190593957901

00:17:29.175 --> 00:17:31.029 have cancer in the lymph nodes than

NOTE Confidence: 0.9190593957901

00:17:31.029 --> 00:17:32.805 what they would have instead is

NOTE Confidence: 0.9190593957901

00:17:32.805 --> 00:17:33.972 just chemotherapy and radiation.

NOTE Confidence: 0.9190593957901

00:17:33.972 --> 00:17:35.604 They save themselves some surgery and

NOTE Confidence: 0.9190593957901

00:17:35.604 --> 00:17:37.524 so that's an example of when we've

NOTE Confidence: 0.9190593957901

00:17:37.524 --> 00:17:39.102 used pretreatment imaging to sort of

NOTE Confidence: 0.9190593957901

00:17:39.110 --> 00:17:41.360 reduce potential extra care or care

NOTE Confidence: 0.9190593957901

00:17:41.360 --> 00:17:44.000 that maybe would not be necessary

NOTE Confidence: 0.9190593957901

00:17:44.000 --> 00:17:46.136 or could have been avoided.

NOTE Confidence: 0.918318212032318

00:17:46.140 --> 00:17:48.908 Do we have datasets with patients who

NOTE Confidence: 0.918318212032318

00:17:48.908 --> 00:17:51.591 were treated in different ways so that

NOTE Confidence: 0.918318212032318

00:17:51.591 --> 00:17:53.910 we can predict given pretreatment data,

NOTE Confidence: 0.918318212032318

00:17:53.910 --> 00:17:55.760 if you got treatment a,

NOTE Confidence: 0.918318212032318

00:17:55.760 --> 00:17:57.610 you will do this well,

NOTE Confidence: 0.918318212032318

00:17:57.610 --> 00:17:59.042 if you got treatment b

NOTE Confidence: 0.918318212032318

00:17:59.042 --> 00:18:02.392 you will do that well and if you got

NOTE Confidence: 0.918318212032318

00:18:02.392 --> 00:18:05.010 treatment c you will do this well.

NOTE Confidence: 0.918318212032318

00:18:05.010 --> 00:18:08.066 And if you got no treatment you would

NOTE Confidence: 0.918318212032318

00:18:08.066 --> 00:18:11.300 do just as well as any of the above.

NOTE Confidence: 0.904884934425354

00:18:11.620 --> 00:18:13.216 Yeah, so that actually kind of

NOTE Confidence: 0.904884934425354

00:18:13.216 --> 00:18:14.624 touches upon something that we're

NOTE Confidence: 0.904884934425354

00:18:14.624 --> 00:18:16.079 really actively exploring and NOTE Confidence: 0.904884934425354

00:18:16.080 --> 00:18:17.760 something that we're very excited about.

NOTE Confidence: 0.904884934425354

00:18:17.760 --> 00:18:19.428 So one thing that

NOTE Confidence: 0.904884934425354

00:18:19.430 --> 00:18:21.110 we've kind of indicated,

NOTE Confidence: 0.904884934425354

00:18:21.110 --> 00:18:22.222 is that these machine

NOTE Confidence: 0.904884934425354

00:18:22.222 --> 00:18:23.334 learning algorithms, these deep

NOTE Confidence: 0.904884934425354

00:18:23.340 --> 00:18:24.452 learning algorithms are extremely

NOTE Confidence: 0.904884934425354

00:18:24.452 --> 00:18:25.564 good at analyzing pictures,

NOTE Confidence: 0.904884934425354

00:18:25.570 --> 00:18:27.937 and so one thing that we've looked at is

NOTE Confidence: 0.904884934425354

00:18:27.937 --> 00:18:30.306 this idea of what we call digital Twins.

NOTE Confidence: 0.904884934425354

00:18:30.310 --> 00:18:31.990 So, based on your pretreatment imaging,

NOTE Confidence: 0.904884934425354

00:18:31.990 --> 00:18:33.106 nothing else no demographics,

NOTE Confidence: 0.904884934425354

00:18:33.106 --> 00:18:34.780 just what your tumor looks like,

NOTE Confidence: 0.904884934425354

00:18:34.780 --> 00:18:36.551 if we could find your digital twin

NOTE Confidence: 0.904884934425354

00:18:36.551 --> 00:18:38.277 or someone who's tumor looks exactly

NOTE Confidence: 0.904884934425354

00:18:38.277 --> 00:18:39.797 like yours or digital family,

NOTE Confidence: 0.904884934425354

00:18:39.800 --> 00:18:41.690 which is maybe a group of five

NOTE Confidence: 0.904884934425354

00:18:41.690 --> 00:18:43.150 people that are like that,

NOTE Confidence: 0.904884934425354

00:18:43.150 --> 00:18:45.446 we can use deep learning to do that.

NOTE Confidence: 0.904884934425354

00:18:45.450 --> 00:18:47.700 And then what we can do is we can

NOTE Confidence: 0.904884934425354

00:18:47.700 --> 00:18:49.908 see, OK among your digital twins,

NOTE Confidence: 0.904884934425354

00:18:49.910 --> 00:18:51.225 or your digital family

NOTE Confidence: 0.904884934425354

00:18:51.225 --> 00:18:53.170 who got one type of treatment,

NOTE Confidence: 0.904884934425354

00:18:53.170 --> 00:18:54.844 this is what their outcome was

NOTE Confidence: 0.904884934425354

00:18:54.844 --> 00:18:56.645 and among your digital family who

NOTE Confidence: 0.904884934425354

00:18:56.645 --> 00:18:58.215 got another type of treatment,

NOTE Confidence: 0.904884934425354

00:18:58.220 --> 00:18:59.996 this is what their outcome was.

NOTE Confidence: 0.904884934425354

00:19:00.000 --> 00:19:00.888 And then they

NOTE Confidence: 0.920491576194763

00:19:00.890 --> 00:19:02.997 can make a more informed decision about

NOTE Confidence: 0.920491576194763

00:19:02.997 --> 00:19:05.047 what they would actually want to do.

NOTE Confidence: 0.920491576194763

00:19:05.050 --> 00:19:07.003 We talk a lot on this

NOTE Confidence: 0.920491576194763

00:19:07.003 --> 00:19:08.689 show about tumor heterogeneity and

NOTE Confidence: 0.920491576194763

00:19:08.689 --> 00:19:10.807 different kinds of cancer and tumor

NOTE Confidence: 0.920491576194763

00:19:10.863 --> 00:19:13.026 biology and all kinds of other things

NOTE Confidence: 0.920491576194763

00:19:13.026 --> 00:19:15.472 that

NOTE Confidence: 0.920491576194763

00:19:15.472 --> 00:19:17.362 seemed to really affect

NOTE Confidence: 0.920491576194763

00:19:17.370 --> 00:19:19.140 biology, they seem to affect outcomes,

NOTE Confidence: 0.920491576194763

00:19:19.140 --> 00:19:21.228 so how is it that you can use that

NOTE Confidence: 0.920491576194763

00:19:21.228 --> 00:19:23.279 data without knowing all of the

NOTE Confidence: 0.920491576194763

00:19:23.279 --> 00:19:25.034 other things?

NOTE Confidence: 0.91830712556839

00:19:25.040 --> 00:19:27.616 I think for one, there's a lot of

NOTE Confidence: 0.91830712556839

00:19:27.616 --> 00:19:29.260 evidence to suggest that deep learning

NOTE Confidence: 0.91830712556839

00:19:29.316 --> 00:19:30.786 algorithms can actually predict changes

NOTE Confidence: 0.91830712556839

00:19:30.786 --> 00:19:33.009 in tumors up to the genomic level,

NOTE Confidence: 0.91830712556839

00:19:33.010 --> 00:19:34.837 so genetic mutations and tumor is based

NOTE Confidence: 0.91830712556839

00:19:34.837 --> 00:19:37.426 off of the pictures because we have to

NOTE Confidence: 0.91830712556839

00:19:37.426 --> 00:19:39.176 appreciate that their really evaluating

NOTE Confidence: 0.91830712556839

00:19:39.230 --> 00:19:41.181 every tumor at a very, very small level,

NOTE Confidence: 0.91830712556839

00:19:41.181 --> 00:19:43.050 every little pixel and each pixel has

NOTE Confidence: 0.91830712556839

00:19:43.106 --> 00:19:44.806 a variety of different intensities,

NOTE Confidence: 0.91830712556839

00:19:44.810 --> 00:19:46.240 and so they're really looking

NOTE Confidence: 0.91830712556839

00:19:46.240 --> 00:19:48.490 at the data in a close level.

NOTE Confidence: 0.91830712556839

00:19:48.490 --> 00:19:51.790 So in lung cancer as well as in brain tumors,

NOTE Confidence: 0.91830712556839

00:19:51.790 --> 00:19:53.566 an also in some lymphomas there's

NOTE Confidence: 0.91830712556839

00:19:53.566 --> 00:19:55.480 been evidence to suggest that deep

NOTE Confidence: 0.91830712556839

00:19:55.480 --> 00:19:57.165 learning algorithms on the diagnostic

NOTE Confidence: 0.91830712556839

00:19:57.165 --> 00:19:59.050 images can predict genomic changes.

NOTE Confidence: 0.91830712556839

00:19:59.050 --> 00:20:00.700 So like driver gene mutations

00:20:02.350 --> 00:20:04.606 that would actually maybe presumably need

NOTE Confidence: 0.91830712556839

00:20:04.606 --> 00:20:06.743 sequencing information for and so then

NOTE Confidence: 0.91830712556839

00:20:06.743 --> 00:20:08.465 that suggests that the pictures actually

NOTE Confidence: 0.91830712556839

00:20:08.465 --> 00:20:10.929 have a lot more information than we think,

NOTE Confidence: 0.91830712556839

00:20:10.930 --> 00:20:13.594 but I do think that you're kind of right

NOTE Confidence: 0.91830712556839

00:20:13.594 --> 00:20:16.716 in one way that maybe it's not just the

NOTE Confidence: 0.91830712556839

00:20:16.716 --> 00:20:18.930 pictures that tell the whole story.

NOTE Confidence: 0.91830712556839

00:20:18.930 --> 00:20:21.132 But the idea is that people

NOTE Confidence: 0.91830712556839

00:20:21.132 --> 00:20:22.600 whose tumors look similar,

NOTE Confidence: 0.91830712556839

00:20:22.600 --> 00:20:24.796 they likely have similar genomic backgrounds.

NOTE Confidence: 0.923692107200623

00:20:24.800 --> 00:20:27.208 I'm still puzzled by

NOTE Confidence: 0.923692107200623

00:20:27.208 --> 00:20:29.470 this whole concept because

NOTE Confidence: 0.923692107200623

00:20:29.470 --> 00:20:32.880 we think about a CT scan or a mammogram,

NOTE Confidence: 0.923692107200623

00:20:32.880 --> 00:20:35.113 or an MRI and each of these

NOTE Confidence: 0.923692107200623

00:20:35.113 --> 00:20:36.910 has its own pitfalls.

NOTE Confidence: 0.923692107200623

00:20:36.910 --> 00:20:39.479 There are false positives on these images,

NOTE Confidence: 0.923692107200623

00:20:39.480 --> 00:20:42.049 there are false negatives on these images,

NOTE Confidence: 0.923692107200623

00:20:42.050 --> 00:20:45.390 so it kind of makes me a little wary to

NOTE Confidence: 0.923692107200623

00:20:45.484 --> 00:20:48.660 put so much faith just in the images.

NOTE Confidence: 0.923692107200623

00:20:48.660 --> 00:20:50.128 How do you explain

NOTE Confidence: 0.92742533882459

00:20:50.130 --> 00:20:51.846 that?

NOTE Confidence: 0.92742533882459

00:20:51.846 --> 00:20:54.218 I guess one of the limitations or one

NOTE Confidence: 0.92742533882459

00:20:54.218 --> 00:20:56.650 of the important caveats to any sort of

NOTE Confidence: 0.92742533882459

00:20:56.650 --> 00:20:58.528 machine learning project is that your

NOTE Confidence: 0.92742533882459

00:20:58.528 --> 00:21:00.948 outcomes are only as good as your data,

NOTE Confidence: 0.92742533882459

00:21:00.950 --> 00:21:03.254 and so if we have a lot of false

NOTE Confidence: 0.92742533882459

00:21:03.254 --> 00:21:05.547 positives in our data set that we have

NOTE Confidence: 0.92742533882459

00:21:05.547 --> 00:21:07.798 not addressed and we haven't identified,

NOTE Confidence: 0.92742533882459

00:21:07.800 --> 00:21:09.746 then I think that it's really important

NOTE Confidence: 0.92742533882459

00:21:09.746 --> 00:21:11.884 that we understand that the machine learning

NOTE Confidence: 0.92742533882459

00:21:11.884 --> 00:21:13.756 algorithm will learn those same errors.

NOTE Confidence: 0.92742533882459

00:21:13.760 --> 00:21:16.442 So similarly any sort of biases that we have,

NOTE Confidence: 0.92742533882459

00:21:16.450 --> 00:21:18.536 maybe we have a bias towards over

NOTE Confidence: 0.92742533882459

00:21:18.536 --> 00:21:20.670 imaging or over

NOTE Confidence: 0.92742533882459

00:21:20.670 --> 00:21:22.150 diagnosing something on an image,

NOTE Confidence: 0.92742533882459

00:21:22.150 --> 00:21:23.980 those similar biases will be promulgated

NOTE Confidence: 0.92742533882459

00:21:23.980 --> 00:21:25.740 through our machine learning algorithms.

NOTE Confidence: 0.92742533882459

00:21:25.740 --> 00:21:27.624 It's actually somewhat of an interesting

NOTE Confidence: 0.92742533882459

00:21:27.624 --> 00:21:29.970 topic in the context of machine learning

NOTE Confidence: 0.92742533882459

00:21:29.970 --> 00:21:31.428 outside of healthcare is that we

NOTE Confidence: 0.92742533882459

00:21:31.428 --> 00:21:33.380 find that a lot of machine learning

NOTE Confidence: 0.92742533882459

00:21:33.380 --> 00:21:35.543 algorithms they mimic the same biases and

NOTE Confidence: 0.92742533882459

00:21:35.595 --> 00:21:36.840 discriminatory abilities

NOTE Confidence: 0.92742533882459

00:21:36.840 --> 00:21:39.745 that people have in regular practice,

NOTE Confidence: 0.92742533882459

00:21:39.750 --> 00:21:41.878 and so a lot of algorithms that

NOTE Confidence: 0.92742533882459

00:21:41.878 --> 00:21:43.989 have been used in law enforcement,

NOTE Confidence: 0.92742533882459

00:21:43.990 --> 00:21:46.132 we find that are actually maybe

NOTE Confidence: 0.92742533882459

00:21:46.132 --> 00:21:48.185 promulgating some of the parts

NOTE Confidence: 0.92742533882459

00:21:48.185 --> 00:21:49.859 of our law enforcement

NOTE Confidence: 0.911179857594626

00:21:49.860 --> 00:21:52.366 that we don't want.

NOTE Confidence: 0.911179857594626

00:21:52.370 --> 00:21:55.274 One way to kind of get around that

NOTE Confidence: 0.911179857594626

00:21:55.274 --> 00:21:58.226 one would think is to use more data,

NOTE Confidence: 0.911179857594626

00:21:58.230 --> 00:21:59.634 not just the images,

NOTE Confidence: 0.911179857594626

00:21:59.634 --> 00:22:02.260 but get down to the genomic level,

NOTE Confidence: 0.911179857594626

00:22:02.260 --> 00:22:04.885 do the biopsy. We've got biopsies on

NOTE Confidence: 0.911179857594626

00:22:04.885 --> 00:22:07.379 most tumors before we ever treat them,

NOTE Confidence: 0.911179857594626

00:22:07.380 --> 00:22:09.576 so look at the pathologic information,

NOTE Confidence: 0.911179857594626

00:22:09.580 --> 00:22:11.410 look at the genomic information.

NOTE Confidence: 0.911179857594626

00:22:11.410 --> 00:22:13.930 We can get a lot of

NOTE Confidence: 0.911179857594626

00:22:13.930 --> 00:22:15.610 sequencing data these days.

NOTE Confidence: 0.911179857594626

00:22:15.610 --> 00:22:16.914 And speaking of which,

NOTE Confidence: 0.911179857594626

00:22:16.914 --> 00:22:19.244 it's really hard to understand what all

NOTE Confidence: 0.911179857594626

00:22:19.244 --> 00:22:21.134 of these different mutations are when

NOTE Confidence: 0.911179857594626

00:22:21.134 --> 00:22:23.429 we think about whole exome sequencing.

NOTE Confidence: 0.911179857594626

00:22:23.430 --> 00:22:25.632 I mean, I would think that

NOTE Confidence: 0.911179857594626

00:22:25.632 --> 00:22:27.100 machine learning might have

NOTE Confidence: 0.911179857594626

00:22:27.170 --> 00:22:28.868 a role to play there too.

00:22:29.210 --> 00:22:31.280 Certainly, I think that machine learning

NOTE Confidence: 0.939351677894592

00:22:31.280 --> 00:22:34.033 in general is probably one of the more

NOTE Confidence: 0.939351677894592

00:22:34.033 --> 00:22:36.001 common approaches to evaluate genomic data.

NOTE Confidence: 0.939351677894592

00:22:36.010 --> 00:22:38.138 Now, because the genome is so complex

NOTE Confidence: 0.939351677894592

00:22:38.138 --> 00:22:41.022 and it's so difficult for us to kind of

NOTE Confidence: 0.939351677894592

00:22:41.022 --> 00:22:43.509 understand that the machine learning

NOTE Confidence: 0.939351677894592

00:22:43.509 --> 00:22:45.819 algorithms are maybe the most common

NOTE Confidence: 0.939351677894592

00:22:45.819 --> 00:22:48.758 ways in which we analyze that sort of

NOTE Confidence: 0.939351677894592

00:22:48.760 --> 00:22:50.330 information now,

NOTE Confidence: 0.939351677894592

00:22:50.330 --> 00:22:52.196 specifically with respect to deep learning,

NOTE Confidence: 0.939351677894592

00:22:52.200 --> 00:22:54.704 which is what our lab is,

NOTE Confidence: 0.939351677894592

00:22:54.710 --> 00:22:55.958 particularly just specific types

NOTE Confidence: 0.939351677894592

00:22:55.958 --> 00:22:57.206 of machine learning methods.

NOTE Confidence: 0.939351677894592

00:22:57.210 --> 00:22:58.775 It's somewhat difficult to evaluate

NOTE Confidence: 0.939351677894592

00:22:58.775 --> 00:23:00.340 genomic information with that data,

NOTE Confidence: 0.939351677894592

00:23:00.340 --> 00:23:02.419 and the reason

NOTE Confidence: 0.939351677894592

00:23:02.419 --> 00:23:04.704 for that is

NOTE Confidence: 0.939351677894592

00:23:04.704 --> 00:23:06.684 because we don't actually have a

NOTE Confidence: 0.939351677894592

00:23:06.751 --> 00:23:09.127 huge data set right now at our

NOTE Confidence: 0.939351677894592

00:23:09.127 --> 00:23:10.600 disposal of genomic information,

NOTE Confidence: 0.939351677894592

00:23:10.600 --> 00:23:13.087 because in order to do some of these

NOTE Confidence: 0.939351677894592

00:23:13.087 --> 00:23:15.368 analysis, we need upwards of 1000 patients,

NOTE Confidence: 0.939351677894592

00:23:15.370 --> 00:23:17.561 and so it's difficult to get 1000

NOTE Confidence: 0.939351677894592

00:23:17.561 --> 00:23:19.300 patients with tumors, images, and

NOTE Confidence: 0.939351677894592

00:23:19.300 --> 00:23:21.300 whole exome sequencing,

NOTE Confidence: 0.939351677894592

00:23:21.300 --> 00:23:22.473 but it's possible,

NOTE Confidence: 0.939351677894592

00:23:22.473 --> 00:23:23.646 and I would

NOTE Confidence: 0.939351677894592

00:23:23.650 --> 00:23:25.954 venture that a place

NOTE Confidence: 0.939351677894592

00:23:25.954 --> 00:23:27.980 like Yale is the place

NOTE Confidence: 0.925964832305908

00:23:27.980 --> 00:23:30.995 that would have that ability to do that or

NOTE Confidence: 0.925964832305908

00:23:30.995 --> 00:23:34.262 or some of these cooperative groups, right?

NOTE Confidence: 0.925964832305908

00:23:34.262 --> 00:23:36.614 For our listeners,

NOTE Confidence: 0.925964832305908

00:23:36.620 --> 00:23:38.978 there are clinical trials that happen

NOTE Confidence: 0.925964832305908

00:23:38.978 --> 00:23:41.336 all across the country, sometimes all

NOTE Confidence: 0.925964832305908

00:23:41.336 --> 00:23:43.694 across the world with cooperative groups.

NOTE Confidence: 0.925964832305908

00:23:43.700 --> 00:23:45.284 These groups of clinicians,

NOTE Confidence: 0.925964832305908

00:23:45.284 --> 00:23:47.660 physicians who are all putting their

NOTE Confidence: 0.925964832305908

00:23:47.725 --> 00:23:50.448 patients on exactly the same clinical trial

NOTE Confidence: 0.925964832305908

00:23:50.450 --> 00:23:52.040 and taking their data,

NOTE Confidence: 0.925964832305908

00:23:52.040 --> 00:23:53.930 putting it in a central repository

NOTE Confidence: 0.925964832305908

00:23:53.930 --> 00:23:56.168 where all of that can be studied.

NOTE Confidence: 0.925964832305908

00:23:56.170 --> 00:23:58.242 Sanjay, I would think that

NOTE Confidence: 0.925964832305908

00:23:58.242 --> 00:24:00.539 would be an ideal place for you

NOTE Confidence: 0.925964832305908

00:24:00.539 --> 00:24:02.850 to get that data.

NOTE Confidence: 0.913027107715607

00:24:02.850 --> 00:24:04.873 Yes, and one effort of our research

NOTE Confidence: 0.913027107715607

00:24:04.873 --> 00:24:06.854 group is actually sort of engaging

NOTE Confidence: 0.913027107715607

00:24:06.854 --> 00:24:08.258 with the cooperative groups.

NOTE Confidence: 0.913027107715607

00:24:08.260 --> 00:24:10.084 There's two that we've

NOTE Confidence: 0.913027107715607

00:24:10.084 --> 00:24:11.760 begun engaging with, the NRG

NOTE Confidence: 0.913027107715607

00:24:11.760 --> 00:24:13.824 which is a large group that has a

NOTE Confidence: 0.913027107715607

00:24:13.824 --> 00:24:16.652 lot of radiation data as well as the

NOTE Confidence: 0.913027107715607

00:24:16.652 --> 00:24:18.839 Southwest Oncology Group also

NOTE Confidence: 0.913027107715607

00:24:18.839 --> 00:24:21.319 known as SWOG in order to sort of develop

NOTE Confidence: 0.913027107715607

00:24:21.320 --> 00:24:22.568 an infrastructure within the

NOTE Confidence: 0.913027107715607

00:24:22.568 --> 00:24:23.816 organization to evaluate machine

NOTE Confidence: 0.913027107715607

00:24:23.816 --> 00:24:25.110 learning techniques and utilized

NOTE Confidence: 0.913027107715607

00:24:25.110 --> 00:24:26.046 machine learning techniques.

NOTE Confidence: 0.913027107715607

00:24:26.050 --> 00:24:27.842 Because a lot of what

NOTE Confidence: 0.913027107715607

00:24:27.842 --> 00:24:29.016 we've designed, these clinical

NOTE Confidence: 0.913027107715607

00:24:29.016 --> 00:24:31.134 trials and these repositories is that

NOTE Confidence: 0.913027107715607

00:24:31.134 --> 00:24:32.660 these cooperative groups,

NOTE Confidence: 0.913027107715607

00:24:32.660 --> 00:24:34.385 their infrastructure wasn't made for

NOTE Confidence: 0.913027107715607

00:24:34.385 --> 00:24:36.494 these sorts of analysis because they

NOTE Confidence: 0.913027107715607

00:24:36.494 --> 00:24:37.814 weren't necessarily thinking that

NOTE Confidence: 0.913027107715607

00:24:37.814 --> 00:24:40.021 this is something that was going to

NOTE Confidence: 0.913027107715607

00:24:40.021 --> 00:24:41.792 come on the horizon.

NOTE Confidence: 0.913027107715607

00:24:41.800 --> 00:24:43.746 And so one thing that we're working

NOTE Confidence: 0.913027107715607

00:24:43.746 --> 00:24:46.061 right now is with SWOG and NRG

NOTE Confidence: 0.913027107715607

00:24:46.061 --> 00:24:48.410 to develop that sort of infrastructure.

NOTE Confidence: 0.913027107715607

00:24:48.410 --> 00:24:50.492 The first process of that is

NOTE Confidence: 0.913027107715607

00:24:50.492 --> 00:24:52.224 developing something that allows us

NOTE Confidence: 0.913027107715607

00:24:52.224 --> 00:24:54.170 to get the imaging data very easily.

NOTE Confidence: 0.913027107715607

00:24:54.170 --> 00:24:56.514 Images are sort of an easy

NOTE Confidence: 0.913027107715607

00:24:56.514 --> 00:24:58.616 method for us to evaluate machine

NOTE Confidence: 0.913027107715607

00:24:58.616 --> 00:25:00.788 learning methods because one,

NOTE Confidence: 0.913027107715607

00:25:00.790 --> 00:25:02.918 it's been shown to be the most effective

NOTE Confidence: 0.913027107715607

00:25:02.918 --> 00:25:05.418 in image analysis across various industries,

NOTE Confidence: 0.913027107715607

00:25:05.420 --> 00:25:06.076 healthcare, technology,

NOTE Confidence: 0.913027107715607

00:25:06.076 --> 00:25:06.404 etc.

NOTE Confidence: 0.913027107715607

00:25:06.404 --> 00:25:08.372 and secondly imaging in healthcare

NOTE Confidence: 0.913027107715607

00:25:08.372 --> 00:25:10.059 has a standardized data format.

NOTE Confidence: 0.913027107715607

00:25:10.060 --> 00:25:11.710 It's a common data model,

NOTE Confidence: 0.913027107715607

00:25:11.710 --> 00:25:14.027 so there's no difficulty about well,

NOTE Confidence: 0.913027107715607

00:25:14.030 --> 00:25:16.016 so and so in California stores

NOTE Confidence: 0.913027107715607

00:25:16.016 --> 00:25:17.340 their data one way,

NOTE Confidence: 0.913027107715607

00:25:17.340 --> 00:25:20.319 and then we store it a different way, etc.

NOTE Confidence: 0.917157232761383

00:25:20.320 --> 00:25:22.637 One thing that you mentioned,

NOTE Confidence: 0.917157232761383

00:25:22.640 --> 00:25:25.631 which I still have to go back to, is

NOTE Confidence: 0.917157232761383

00:25:25.631 --> 00:25:27.035 you said that you're

NOTE Confidence: 0.917157232761383

00:25:27.035 --> 00:25:28.439 interested in deep learning,

NOTE Confidence: 0.917157232761383

00:25:28.440 --> 00:25:31.272 which is a type of machine learning that

NOTE Confidence: 0.917157232761383

00:25:31.272 --> 00:25:33.570 is particularly well suited to imaging.

NOTE Confidence: 0.917157232761383

00:25:33.570 --> 00:25:35.280 Tell us the difference between

NOTE Confidence: 0.917157232761383

00:25:35.280 --> 00:25:36.990 deep learning and machine learning.

NOTE Confidence: 0.917157232761383

00:25:36.990 --> 00:25:38.010 Yeah, that's a

NOTE Confidence: 0.930824279785156

00:25:38.010 --> 00:25:39.720 good question,

NOTE Confidence: 0.930824279785156

00:25:39.720 --> 00:25:41.430 and the words,

NOTE Confidence: 0.930824279785156

00:25:41.430 --> 00:25:42.114 artificial intelligence,

NOTE Confidence: 0.930824279785156

00:25:42.114 --> 00:25:42.798 machine learning,

NOTE Confidence: 0.930824279785156

00:25:42.798 --> 00:25:43.824 and deep learning.

NOTE Confidence: 0.930824279785156

00:25:43.830 --> 00:25:45.876 sort of get thrown around together,

NOTE Confidence: 0.930824279785156

00:25:45.880 --> 00:25:48.267 and it's difficult to parse them out.

NOTE Confidence: 0.930824279785156

00:25:48.270 --> 00:25:49.638 I think that machine

NOTE Confidence: 0.930824279785156

00:25:49.638 --> 00:25:51.006 learning is a broad

NOTE Confidence: 0.930824279785156

00:25:51.010 --> 00:25:52.720 discipline of various types of

NOTE Confidence: 0.930824279785156

00:25:52.720 --> 00:25:54.430 mathematical techniques to model data.

NOTE Confidence: 0.930824279785156

00:25:54.430 --> 00:25:56.190 Deep learning is just one

NOTE Confidence: 0.930824279785156

00:25:56.190 --> 00:25:57.246 of those techniques.

NOTE Confidence: 0.930824279785156

00:25:57.250 --> 00:25:58.735 Now the difference between deep

NOTE Confidence: 0.930824279785156

00:25:58.735 --> 00:26:00.220 learning and other traditional machine

NOTE Confidence: 0.930824279785156

00:26:00.266 --> 00:26:01.916 learning techniques is that other

NOTE Confidence: 0.930824279785156

00:26:01.916 --> 00:26:03.236 machine learning techniques require,

NOTE Confidence: 0.930824279785156

00:26:03.240 --> 00:26:05.094 you know inputs that are called

NOTE Confidence: 0.930824279785156

00:26:05.094 --> 00:26:07.335 features and so they can only handle

NOTE Confidence: 0.930824279785156

00:26:07.335 --> 00:26:09.540 data that comes in a featured format.

NOTE Confidence: 0.930824279785156

00:26:09.540 --> 00:26:11.358 So sort of predictor variables that

NOTE Confidence: 0.930824279785156

00:26:11.358 --> 00:26:12.570 you're interested in demographic

NOTE Confidence: 0.930824279785156

00:26:12.621 --> 00:26:13.913 variables or variables from

NOTE Confidence: 0.930824279785156

00:26:13.913 --> 00:26:15.205 the electronic medical record.

NOTE Confidence: 0.930824279785156

00:26:15.210 --> 00:26:16.606 Deep learning is particularly

NOTE Confidence: 0.930824279785156

00:26:16.606 --> 00:26:18.700 unique in that it doesn't actually

NOTE Confidence: 0.930824279785156

00:26:18.758 --> 00:26:20.557 require data at all from a human.

NOTE Confidence: 0.930824279785156

00:26:20.560 --> 00:26:22.135 It doesn't require any sort

NOTE Confidence: 0.930824279785156

00:26:22.135 --> 00:26:23.080 of human interaction.

NOTE Confidence: 0.930824279785156

00:26:23.080 --> 00:26:24.910 It can learn those features on

NOTE Confidence: 0.930824279785156

00:26:24.910 --> 00:26:27.469 its own as long as it has access

NOTE Confidence: 0.930824279785156

00:26:27.469 --> 00:26:29.377 to what they called the sensor.

NOTE Confidence: 0.930824279785156

00:26:29.380 --> 00:26:32.260 So where the data is generated and as data

NOTE Confidence: 0.930824279785156

00:26:32.260 --> 00:26:34.240 is being generated in real time,

NOTE Confidence: 0.930824279785156

00:26:34.240 --> 00:26:36.610 deep learning algorithms can analyze it,

NOTE Confidence: 0.930824279785156

00:26:36.610 --> 00:26:38.186 identify those features that

NOTE Confidence: 0.930824279785156

00:26:38.186 --> 00:26:39.368 are very important,

NOTE Confidence: 0.930824279785156

00:26:39.370 --> 00:26:42.530 so those predictors are important and then

NOTE Confidence: 0.905940711498261

00:26:42.530 --> 00:26:44.110 create predictions.

NOTE Confidence: 0.905940711498261

00:26:44.110 --> 00:26:46.085 How exactly does this happen?

NOTE Confidence: 0.905940711498261

00:26:46.090 --> 00:26:48.060 Somebody's gotta

NOTE Confidence: 0.905940711498261

00:26:48.060 --> 00:26:50.430 program this thing right?

NOTE Confidence: 0.905940711498261

00:26:50.430 --> 00:26:51.970 Yes, it is programmed,

NOTE Confidence: 0.905940711498261

00:26:51.970 --> 00:26:54.773 typically in Python and so the way

NOTE Confidence: 0.905940711498261

00:26:54.773 --> 00:26:57.257 that the process works for developing

NOTE Confidence: 0.905940711498261

00:26:57.257 --> 00:26:59.692 a deep learning algorithm is first

NOTE Confidence: 0.905940711498261

00:26:59.692 --> 00:27:02.289 you have a set of training data.

NOTE Confidence: 0.905940711498261

00:27:02.290 --> 00:27:04.257 And the associated labels to that data.

NOTE Confidence: 0.905940711498261

00:27:04.260 --> 00:27:05.670 So you already have data

NOTE Confidence: 0.905940711498261

00:27:05.670 --> 00:27:07.080 with outcomes that you know,

NOTE Confidence: 0.905940711498261

00:27:07.080 --> 00:27:08.826 and that's the process you're

NOTE Confidence: 0.905940711498261

00:27:08.826 --> 00:27:10.836 trying to predict and then what you

NOTE Confidence: 0.905940711498261

00:27:10.836 --> 00:27:12.684 do is you design your deep learning

NOTE Confidence: 0.905940711498261

00:27:12.738 --> 00:27:14.138 algorithm using a complex series

NOTE Confidence: 0.905940711498261

00:27:14.138 --> 00:27:16.026 of what they call neural networks

NOTE Confidence: 0.905940711498261

00:27:16.026 --> 00:27:18.337 and what we do is we kind of train

NOTE Confidence: 0.905940711498261

00:27:18.337 --> 00:27:20.500 the algorithm by looking at each of

NOTE Confidence: 0.905940711498261

00:27:20.500 --> 00:27:22.308 those training data set labels to

NOTE Confidence: 0.905940711498261

00:27:22.308 --> 00:27:24.264 identify sort of patterns in the data.

NOTE Confidence: 0.905940711498261

00:27:24.264 --> 00:27:26.112 And it takes a significant amount of

NOTE Confidence: 0.905940711498261

00:27:26.112 --> 00:27:28.155 time and a huge amount of computational

NOTE Confidence: 0.905940711498261

00:27:28.155 --> 00:27:29.930 resources in order to do that.

NOTE Confidence: 0.905940711498261

00:27:29.930 --> 00:27:31.310 So deep learning algorithms that

NOTE Confidence: 0.905940711498261

00:27:31.310 --> 00:27:33.054 we developed in our lab oftentimes

NOTE Confidence: 0.905940711498261

00:27:33.054 --> 00:27:34.278 take weeks to train.

NOTE Confidence: 0.905940711498261

00:27:34.280 --> 00:27:36.746 And so we just let it run all week

NOTE Confidence: 0.905940711498261

00:27:36.746 --> 00:27:39.265 and let it run every single time

NOTE Confidence: 0.905940711498261

00:27:39.265 --> 00:27:41.594 and every piece of information

NOTE Confidence: 0.905940711498261

00:27:41.594 --> 00:27:42.806 that it gets

NOTE Confidence: 0.905940711498261

00:27:42.810 --> 00:27:44.170 it looks at the

NOTE Confidence: 0.912006855010986

00:27:44.170 --> 00:27:46.228 outcome and it tries to learn

NOTE Confidence: 0.912006855010986

00:27:46.228 --> 00:27:48.260 a little bit more from it.

NOTE Confidence: 0.912006855010986

00:27:48.260 --> 00:27:50.647 And so presumably you

NOTE Confidence: 0.912006855010986

00:27:50.647 --> 00:27:52.962 design this algorithm, you let it run,

NOTE Confidence: 0.912006855010986

00:27:52.962 --> 00:27:56.030 the machine tries to learn what it can to

NOTE Confidence: 0.912006855010986

00:27:56.030 --> 00:27:58.487 try and improve its prediction each time,

NOTE Confidence: 0.912006855010986

00:27:58.490 --> 00:28:01.560 and then you test it on a separate set.

NOTE Confidence: 0.912006855010986

00:28:01.560 --> 00:28:04.666 Yes, exactly and in order for

NOTE Confidence: 0.912006855010986

00:28:04.666 --> 00:28:07.898 this to be utilized because

NOTE Confidence: 0.912006855010986

00:28:07.900 --> 00:28:11.127 I'm thinking about how this can be

NOTE Confidence: 0.912006855010986

00:28:11.127 --> 00:28:14.582 utilized in the clinic in terms of you

NOTE Confidence: 0.912006855010986

00:28:14.582 --> 00:28:18.410 know one day you may walk into a clinic,

NOTE Confidence: 0.912006855010986

00:28:18.410 --> 00:28:21.476 have your CT scan, have your biopsy,

NOTE Confidence: 0.912006855010986

00:28:21.480 --> 00:28:24.192 fill out some paperwork on your

NOTE Confidence: 0.912006855010986

00:28:24.192 --> 00:28:27.349 demographics and I can imagine a time

NOTE Confidence: 0.912006855010986

00:28:27.349 --> 00:28:30.808 when all of that information is put into

NOTE Confidence: 0.912006855010986

00:28:30.808 --> 00:28:33.888 a computer or stored in the computer.

NOTE Confidence: 0.912006855010986

00:28:33.890 --> 00:28:36.902 This algorithm runs in the background

NOTE Confidence: 0.912006855010986

00:28:36.902 --> 00:28:40.309 and spits out to the clinician,

NOTE Confidence: 0.912006855010986

00:28:40.310 --> 00:28:42.990 this patient's prognosis is X.

NOTE Confidence: 0.912006855010986

00:28:42.990 --> 00:28:47.814 The ideal treatment out of A, B&C is B

NOTE Confidence: 0.912006855010986

00:28:47.820 --> 00:28:49.878 and that kind of thing.

00:28:50.170 --> 00:28:51.350 I could see that happening

NOTE Confidence: 0.916711211204529

00:28:51.350 --> 00:28:52.916 in the future. I think we're

NOTE Confidence: 0.916711211204529

00:28:52.916 --> 00:28:55.004 a little bit far away from

NOTE Confidence: 0.916711211204529

00:28:55.004 --> 00:28:56.639 complete automation in that way,

NOTE Confidence: 0.916711211204529

00:28:56.640 --> 00:28:58.464 and partly because I think that

NOTE Confidence: 0.916711211204529

00:28:58.464 --> 00:29:01.078 we don't have a good hold on

NOTE Confidence: 0.916711211204529

00:29:01.078 --> 00:29:03.371 the data that we think is the most

NOTE Confidence: 0.916711211204529

00:29:03.371 --> 00:29:05.457 important and we don't have a good

NOTE Confidence: 0.916711211204529

00:29:05.460 --> 00:29:07.218 way of storing all that information,

NOTE Confidence: 0.916711211204529

00:29:07.220 --> 00:29:09.695 but I think that it's not something that I

NOTE Confidence: 0.916711211204529

00:29:09.695 --> 00:29:12.220 would be surprised we're doing in 5-10 years.

NOTE Confidence: 0.939991235733032

00:29:12.810 --> 00:29:15.449 Doctor Sanjay Aneja is an assistant

NOTE Confidence: 0.939991235733032

00:29:15.449 --> 00:29:16.994 professor of therapeutic radiology

NOTE Confidence: 0.939991235733032

00:29:16.994 --> 00:29:19.094 at the Yale School of Medicine.

NOTE Confidence: 0.939991235733032

00:29:19.100 --> 00:29:20.584 If you have questions,

NOTE Confidence: 0.939991235733032

00:29:20.584 --> 00:29:22.068 the address is canceranswers@yale.edu

NOTE Confidence: 0.939991235733032

00:29:22.068 --> 00:29:24.118 and past editions of the program

NOTE Confidence: 0.939991235733032

00:29:24.118 --> 00:29:25.990 are available in audio and written

NOTE Confidence: 0.939991235733032

00:29:26.048 --> 00:29:27.608 form at Yalecancercenter.org.

NOTE Confidence: 0.939991235733032

00:29:27.610 --> 00:29:30.090 We hope you'll join us next week to

NOTE Confidence: 0.939991235733032

00:29:30.090 --> 00:29:32.499 learn more about the fight against

NOTE Confidence: 0.939991235733032

00:29:32.499 --> 00:29:35.016 cancer here on Connecticut public radio.