WEBVTT

1 00:00:00.000 --> 00:00:04.129 Support for Connecticut Public Radio comes from AstraZeneca,

2 00:00:04.129 --> 00:00:10.910 a biopharmaceutical business that is pushing the boundaries of science to deliver new cancer

3 00:00:10.910 --> 00:00:14.960 medicine. More information at astrazeneca-us.com.

4 00:00:14.960 --> 00:00:20.390 Welcome to Yale Cancer Answers with doctor Anees Chagpar.

5 00:00:20.390 --> 00:00:30.890 Yale Cancer Answers features the latest information on cancer care by welcoming oncologists and specialists who are on the forefront of the battle to fight cancer. This week,

6 00:00:30.890 --> 00:00:34.509 it's a conversation about machine learning and prostate cancer treatment

7 00:00:34.509 --> 00:00:43.560 with doctor John Onofrey. Doctor Onofrey is an Assistant Professor of Radiology and Biomedical Imaging and of Urology at Yale School of Medicine.

9 00:00:47.869 --> 00:00:48.229 John, let's start

10 00:00:48.229 --> 00:00:54.619 off by having you tell us a little bit about yourself and what exactly you do.

11 00:00:54.619 --> 00:00:57.810 I have a background in computer science,

12 00:00:57.810 --> 00:01:05.620 so I actually spent four years working as a software engineer in the defense industry before coming back to get my PhD,

13 00:01:05.620 --> 00:01:07.750 which I actually did here at Yale.

14 00:01:07.750 --> 00:01:18.060 In that time I became interested in medical image processing and part of that that became a driving factor was the use of machine learning and artificial intelligence to

15 00:01:18.060 --> 00:01:20.359 create solutions for image analysis problems,

16 00:01:20.359 --> 00:01:23.040 and particularly those applied to radiology.

18 00:01:23.420 --> 00:01:25.719 All of that sounded really cool,

19 00:01:25.719 --> 00:01:31.079 but you're kind of losing me in terms of what exactly you are talking about.

20 00:01:31.079 --> 00:01:42.569 People go and they get X Rays and CT scans and ultrasounds and those kinds of things as diagnostic tests and some of us may have heard

21 00:01:42.569 --> 00:01:46.019 a little bit about artificial intelligence and machine learning,

22 00:01:46.019 --> 00:01:49.469 but it seems to be this amorphis concept like

23 00:01:49.469 --> 00:01:54.980 are machines actually going to learn how to do the job of humans?

24 00:01:54.980 --> 00:01:58.280 Are they going to take over what we do?

25 00:01:58.280 --> 00:02:05.620 Put that whole concept together for me and explain a little bit about what exactly is the marriage between those two things.

26 00:02:05.620 --> 00:02:10.759 Artificial intelligence and machine learning really is a very broad concept,

27 00:02:10.759 --> 00:02:18.479 and it's especially a very broad range in terms of medical diagnosis or any kind of medical decision making.

28 00:02:18.479 --> 00:02:20.400 A lot of problems involved though.

29 00:02:20.400 --> 00:02:23.599 What's something that the computer can help a clinician do?

30 00:02:23.599 --> 00:02:32.240 Is there a task that the computer can aid them in some way so that they can do their job either better or more efficiently?

31 00:02:32.240 --> 00:02:34.479 Especially in a imaging,

32 00:02:34.479 --> 00:02:39.919 the most basic task is, well can I identify some part of an image that is of interest.

33 00:02:39.919 --> 00:02:42.159 So for example in prostate cancer care,

34 00:02:42.159 --> 00:02:47.280 one of the preliminary steps in any analysis is just to identify the prostate gland itself,

35 00:02:47.280 --> 00:02:49.250 and it turns out a machine.

36 00:02:49.250 --> 00:02:50.879 is able to do that

37 00:02:50.879 --> 00:02:59.650 if you have someone to teach it and that data is very important and that data comes from these radiologists that are available at our institution,

38 00:02:59.650 --> 00:03:01.930 so it's really what data goes in,

39 00:03:01.930 --> 00:03:05.180 the machine learns what these radiologists do,

40 00:03:05.180 --> 00:03:14.280 hopefully they can do it as well and spit out an answer and try to do in an automated fashion and that way you can hopefully aid this clinician

41 00:03:14.280 --> 00:03:20.639 with their job.

42 00:03:20.639 --> 00:03:25.949 So you have an image like a CT scan, the prostate is a part that we can find on the see CT scan.

43 00:03:25.949 --> 00:03:34.090 And so if the radiologist, who are used to looking at CT scans, can teach the computer what a prostate gland looks like,

44 00:03:34.090 --> 00:03:36.219 then the computer can identify it.

45 00:03:36.219 --> 00:03:42.939 But then the question becomes, the radiologist is more than looking at where the prostate gland is,

46 00:03:42.939 --> 00:03:47.539 they are the ones who say is there something wrong with the prostate?

47 00:03:47.539 --> 00:03:50.020 Is there a nodule in the prostate,

48 00:03:50.020 --> 00:03:52.699 is there a cancer lurking in that prostate?

49 00:03:52.699 --> 00:03:55.300 Can the computers help us with that too?

50 00:03:55.300 --> 00:03:55.629 Absolutely.

51 00:03:55.629 --> 00:03:59.199 Just to clarify though. Actually in the prostate radiology world,

52 00:03:59.199 --> 00:04:02.780 actually most of the imaging is done with magnetic resonance imaging,

53 00:04:02.780 --> 00:04:08.949 so that just gives a richer sense of that issue that's within the prostate compared to something like CT.

54 00:04:08.949 --> 00:04:10.900 But yes, to answer your question,

55 00:04:10.900 --> 00:04:13.500 so whenever a radiologist looks at this image,

56 00:04:13.500 --> 00:04:18.050 they have years and years of training that goes into what to look for.

57 00:04:18.050 --> 00:04:22.350 So not only are they looking at just the shape of the prostate,

58 00:04:22.350 --> 00:04:30.040 but they make a diagnosis on what they think is suspected cancer and those manifest in different ways in this image.

59 00:04:30.040 --> 00:04:32.230 So they look for different patterns,

60 00:04:32.230 --> 00:04:36.620 different textures, and it all comes with years and years of training.

61 00:04:36.620 --> 00:04:43.579 So essentially what we do is we have that radiologist with their pre annotated results.

62 00:04:43.579 --> 00:04:47.240 So they mark up this image somehow with their tool.

63 00:04:47.240 --> 00:04:53.389 They'll say, well, I think this has some level of prostate cancer risk or some assessment.

64 00:04:53.389 --> 00:04:55.589 And then we can take that data,

65 00:04:55.589 --> 00:04:58.410 both the original image and what their labeling is,

66 00:04:58.410 --> 00:04:59.980 put it into an algorithm,

67 00:04:59.980 --> 00:05:03.750 and then hopefully that algorithm can learn to do a similar thing.

68 00:05:03.750 --> 00:05:05.319 Now the goal is,

69 00:05:05.319 --> 00:05:11.290 can you actually achieve some kind of performance that applies to all the datasets that you haven't seen?

70 00:05:11.290 --> 00:05:13.490 That's a real challenge in artificial intelligence.

71 00:05:13.490 --> 00:05:19.449 Can you get something that you've never seen before and that's one of the big questions that we have.

72 00:05:19.449 --> 00:05:23.529 So what we're really trying to distill is all the knowledge within this model.

73 00:05:23.529 --> 00:05:26.100 Just think of it as a black box.

74 00:05:26.100 --> 00:05:30.329 Can we capture what these radiologists have taught within this black box and so

75 00:05:30.329 --> 00:05:35.160 essentially the question is, can one day the computer take over the job of the radiologist?

76 00:05:35.160 --> 00:05:37.269 I don't think so. That seems to

77 00:05:37.269 --> 00:05:46.329 be everyone's fear. I look at it more as it could be a helpful assistant and aid like a clinical diagnostic tool that they can leverage to improve their own

78 00:05:46.329 --> 00:05:55.389 level of care and also see a very big point of this could be at Yale were very fortunate to have lots of experts doing this kind of imaging,

79 00:05:55.389 --> 00:05:57.300 one of the main challenges is,

80 00:05:57.300 --> 00:06:00.850 what if you have someone who is not an expert trained in this?

81 00:06:00.850 --> 00:06:03.389 Will they perform as well as the expert?

82 00:06:03.389 --> 00:06:07.209 Most likely no. But if you're able to give them this tool,

83 00:06:07.209 --> 00:06:11.660 can we bring that more novice reader up to the level of the expert?

84 00:06:11.660 --> 00:06:19.290 And can you disseminate this technology down into lower centers of care that it could be really impactful to patient health across the population?

87 00:06:21.519 --> 00:06:27.240 For example, if you're in the community and you don't have one of these experienced radiologists,

88 00:06:27.240 --> 00:06:29.149 maybe you have a general radiologist.

89 00:06:29.149 --> 00:06:34.560 The computer might be able to show them a spot that maybe they should be more worried

90 00:06:34.560 --> 00:06:37.689 about.

91 00:06:37.689 --> 00:06:44.389 This machine learning could highlight an area of interest and you never want to say that that area of interest is definitely cancer,

92 00:06:44.389 --> 00:06:47.459 but what we want to do is point it out to the radiologist.

93 00:06:47.459 --> 00:06:50.800 Make them aware, maybe it was something that they would have missed,

94 00:06:50.800 --> 00:06:52.759 that they would have not seen otherwise.

95 00:06:52.759 --> 00:06:55.829 But if they take a second look because of this algorithm,

96 00:06:55.829 --> 00:06:57.779 then that means we've done our job,

97 00:06:57.779 --> 00:07:02.240 especially if it leads to that was actually something that they should have been looking at,

98 00:07:02.240 --> 00:07:03.639 and they just happened to

99 00:07:03.639 --> 00:07:05.600 overlook it. And I think that's

100 00:07:05.600 --> 00:07:11.019 possible because humans are human and suffer from fatigue or whatever else absolutely,

101 00:07:11.019 --> 00:07:13.899 so that's usually the next step after diagnosis.

102 00:07:13.899 --> 00:07:18.959 Once you have the image and you see something that looks a little funny,

103 00:07:18.959 --> 00:07:21.120 the next step is a biopsy.

104 00:07:21.120 --> 00:07:25.089 Will artificial intelligence and machine learning help us in that?

105 00:07:25.089 --> 00:07:26.180 So that's actually

106 00:07:26.180 --> 00:07:29.430 one area of research that I've been involved in,

107 00:07:29.430 --> 00:07:33.040 how to improve the targeting of that biopsy.

108 00:07:33.040 --> 00:07:36.019 So when a patient goes for a biopsy,

109 00:07:36.019 --> 00:07:37.899 they do so under ultrasound guidance,

110 00:07:37.899 --> 00:07:41.360 so a urologist has the ability to see what their targeting,

111 00:07:41.360 --> 00:07:46.379 but they aren't able to discern what is a cancerous lesion or not of the prostate.

112 00:07:46.379 --> 00:07:50.149 However, that lesion is able to be discerned on the MRI.

113 00:07:50.149 --> 00:07:55.800 The problem then becomes how do you map your target in your MRI image to your ultrasound,

114 00:07:55.800 --> 00:08:03.649 and that's where we came in to develop a model that could actually predict the way that the prostate would change during the two procedures,

115 00:08:03.649 --> 00:08:06.220 so it provided a way to hopefully more

116 00:08:06.220 --> 00:08:11.180 accurately target these so by imagining it like having a bullseye,

117 00:08:11.180 --> 00:08:17.730 we want to show where exactly that urologist should aim their biopsy needle.

118 00:08:17.730 --> 00:08:18.069 So how do you do that exactly?

120 00:08:20.089 --> 00:08:30.199 Because we've had urologists on the show before and they've talked about how they can see things on the MRI and when they go to ultrasound they really

121 00:08:30.199 --> 00:08:33.470 can't. And so sometimes these biopsies are almost,

122 00:08:33.470 --> 00:08:35.990 I don't want to say random,

123 00:08:35.990 --> 00:08:39.350 but almost because you can't necessarily correlate it,

124 00:08:39.350 --> 00:08:43.549 especially if there's no palpable lesion that you can feel,

125 00:08:43.549 --> 00:08:51.110 so how does the computer take an image on one modality is completely different?

126 00:08:51.110 --> 00:08:55.730 They look nothing alike either and translate it into another modality.

127 00:08:55.730 --> 00:08:59.559 I mean,

128 00:08:59.559 --> 00:09:01.149 an ultrasound is completely different.

129 00:09:01.149 --> 00:09:02.730 How do you do that?

130 00:09:02.730 --> 00:09:03.679 We actually are

131 00:09:03.679 --> 00:09:06.220 able to leverage human intelligence in this case,

132 00:09:06.220 --> 00:09:11.919 so both the radiologist and the urologist provide an initial guess about where the prostate gland is itself.

133 00:09:11.919 --> 00:09:13.830 So first on the radiology side,

134 00:09:13.830 --> 00:09:15.409 a radiologist will actually contour,

135 00:09:15.409 --> 00:09:17.950 we call it segmentation of the prostate gland,

136 00:09:17.950 --> 00:09:21.120 and that takes a few minutes to do, and again,

137 00:09:21.120 --> 00:09:24.600 this gets back to something that I was talking about earlier.

138 00:09:24.600 --> 00:09:27.460 Can you have a computer program do that automatically?

139 00:09:27.460 --> 00:09:31.850 So there's one way that we can improve the efficiency of the workflow.

140 00:09:31.850 --> 00:09:34.509 But right now we manually have to do it,

141 00:09:34.509 --> 00:09:36.279 because that's what we rely upon,

142 00:09:36.279 --> 00:09:39.820 and the urologist will actually do the same thing in the ultrasound.

143 00:09:39.820 --> 00:09:43.059 While they're doing the procedure before it starts for the biopsy,

144 00:09:43.059 --> 00:09:47.490 they will contour this ultrasound and they will find out where the prostate gland is.

145 00:09:47.490 --> 00:09:51.029 So now we have two shapes of what the prostate looks like,

146 00:09:51.029 --> 00:09:53.980 one in the MR imaging and one in the ultrasound.

147 00:09:53.980 --> 00:09:58.399 So now now that we have these surface, these shapes were able to co register,

148 00:09:58.399 --> 00:09:59.879 we call this image fusion,

149 00:09:59.879 --> 00:10:02.799 we actually bring the two into alignment.

150 00:10:02.799 --> 00:10:05.049 And by using these models instead,

151 00:10:05.049 --> 00:10:07.309 these surfaces, instead of the image in itself.

152 00:10:07.309 --> 00:10:13.750 That's how we kind of get away with the very different appearances of these images in the two different imaging

153 00:10:13.750 --> 00:10:20.820 modalities.

154 00:10:20.820 --> 00:10:27.700 I get the fact that you contour it out and you say here is the prostate in this ball.

155 00:10:27.700 --> 00:10:31.830 And here is the prostate in this other ball on the ultrasound.

156 00:10:31.830 --> 00:10:38.019 But to put them together because then ultimately you have to feed that information to the urologist,

157 00:10:38.019 --> 00:10:44.210 not only to say, you know that ball that you were thinking was the prostate on the ultrasound,

158 00:10:44.210 --> 00:10:49.970 well here it is. How it looks on the MR and

159 00:10:49.970 --> 00:10:54.309 oh, by the way, the lesion that we're going after is here,

160 00:10:54.309 --> 00:10:57.210 which you can't really see on the ultrasound,

161 00:10:57.210 --> 00:11:04.450 but you're going to have to trust us that it's kinda here in this fused image that you can't really see.

162 00:11:04.450 --> 00:11:07.700 Correct. What we do is basically that fusion,

163 00:11:07.700 --> 00:11:09.509 like I said before,

164 00:11:09.509 --> 00:11:15.370 it provides a target so that target is displayed in real time on the ultrasound image.

165 00:11:15.370 --> 00:11:17.210 So when the urologist is performing

166 00:11:17.210 --> 00:11:24.549 the procedure they look at the ultrasound image and the beauty of ultrasound is that it is in real time.

167 00:11:24.549 --> 00:11:28.830 So what you see is what you are looking at currently in real time,

168 00:11:28.830 --> 00:11:35.259 and so the software is actually able to transform and fuse that lesion on to that image in real time.

169 00:11:35.259 --> 00:11:38.009 So then the urologist is able to target it.

170 00:11:38.009 --> 00:11:40.159 That's where they aim the biopsy needle,

171 00:11:40.159 --> 00:11:46.019 and so the particular device that's here at Yale actually has a mechanical arm that stabilizes the biopsy procedure.

172 00:11:46.019 --> 00:11:47.440 And so it's a

173 00:11:47.440 --> 00:11:51.129 Iknown trajectory on where that biopsy needle is going to go,

174 00:11:51.129 --> 00:11:53.690 and so able to not only target the lesion,

175 00:11:53.690 --> 00:11:56.240 but also records where that biopsy sample was performed,

176 00:11:56.240 --> 00:12:00.500 and so that actually gets into the downstream effects of when that goes to pathology.

177 00:12:00.500 --> 00:12:01.639 Did you actually hit

178 00:12:01.639 --> 00:12:04.480 that lesion which was going to be my next question?

179 00:12:04.480 --> 00:12:08.740 Because you can tell me that the target is at Point X on the ultrasound,

180 00:12:08.740 --> 00:12:11.580 but if I can't see Point X on the ultrasound,

181 00:12:11.580 --> 00:12:13.279 I'm kind of taking your word

182 00:12:13.279 --> 00:12:17.344 for it. You are putting your trust entirely in the fusion algorithm itself,

183 00:12:17.395 --> 00:12:25.220 right? Which is particularly interesting because the segmentation or the outlining of that gland on the ultrasound is extremely challenging.

184 00:12:25.220 --> 00:12:29.000 Urologist have a very difficult time and it's not against them.

185 00:12:29.000 --> 00:12:35.929 I mean they have years of training and you ask the same urologist to do the same person again,

186 00:12:35.929 --> 00:12:44.120 you'll get a different answer and that's actually where the innovation and the research that we've been doing here at Yale comes in.

187 00:12:44.120 --> 00:12:46.330 Can we handle these kinds of mistakes?

188 00:12:46.330 --> 00:12:50.740 These errors that are going to happen no matter what.

189 00:12:50.740 --> 00:12:52.940 Can we make a more robust fusion

190 00:12:52.940 --> 00:13:00.480 that is less sensitive to these kinds of problems and so you have the variability in the urologist outlining the prostate

191 00:13:00.480 --> 00:13:10.860 and then you have the fact that they can't see the lesion and you give them a target and you tell them aim here and the biopsy is taken there.

192 00:13:10.860 --> 00:13:13.629 Have you looked at how often you're right?

193 00:13:14.169 --> 00:13:15.909 We're actually quantifying that right now,

196 00:13:18.230 --> 00:13:22.000 Not only pathology, but what if on the MR are were wrong,

197 00:13:22.000 --> 00:13:27.220 right? So to go back and look at the MR and say I did the biopsy here,

198 00:13:27.220 --> 00:13:30.120 was it actually the place where we meant to target?

199 00:13:30.120 --> 00:13:32.440 Because we can see it on the MR.

200 00:13:32.440 --> 00:13:37.080 We actually do that in tumor board when we get everybody together in a room.

201 00:13:37.080 --> 00:13:39.690 We get the radiologist. We get the pathologist,

202 00:13:39.690 --> 00:13:44.419 altogether and what we do is we look at what cases we possibly missed.

203 00:13:44.419 --> 00:13:47.820 And that's a very useful thing.

204 00:13:47.820 --> 00:13:48.250 So

205 00:13:48.250 --> 00:13:50.799 we're actually going backwards from results.

206 00:13:50.799 --> 00:13:56.320 There's a lot more to talk about about in AI and prostate cancer,

207 00:13:56.320 --> 00:13:57.600 right after we

208 00:13:57.600 --> 00:14:10.350 take a short break for a medical minute. Support for Connecticut Public Radio comes from AstraZeneca working side by side with leading scientists to better understand how complex data

209 00:14:10.350 --> 00:14:13.830 can be converted into innovative treatments. More information at astrazeneca-us.com.

210 00:14:13.830 --> 00:14:17.139 This is a medical minute about pancreatic cancer,

211 00:14:17.139 --> 00:14:21.279 which represents about 3% of all cancers in the US,

212 00:14:21.279 --> 00:14:23.350 and about 7% of cancer deaths.

213 00:14:23.350 --> 00:14:33.289 Clinical trials are currently being offered at federally designated comprehensive Cancer Centers for the treatment of advanced stage and metastatic pancreatic cancer using chemotherapy and

214 00:14:33.289 --> 00:14:36.190 other novel therapies, FOLFIRINOX

215 00:14:36.190 --> 00:14:43.220 a combination of five different chemotherapies is the latest advances in the treatment of metastatic pancreatic cancer,

216 00:14:43.220 --> 00:14:48.309 and research continues. It centers around the work looking into targeted therapies.

217 00:14:48.309 --> 00:14:58.149 and a recently discovered marker hENT-1. This has been a medical minute brought to you as a public service by Yale Cancer Center.

218 00:14:58.149 --> 00:15:03.659 More information is available at yalecancercenter.org, you're listening to Connecticut Public Radio.

219 00:15:03.659 --> 00:15:04.470 Now John,

220 00:15:04.470 --> 00:15:13.330 right before the break we were saying that the urologist really puts their trust in this targeting device,

221 00:15:13.330 --> 00:15:15.750 because they can't see the lesion.

222 00:15:15.750 --> 00:15:18.970 The lesion shows up on the MR

223 00:15:18.970 --> 00:15:21.799 but they're doing the biopsy under ultrasound,

224 00:15:21.799 --> 00:15:23.809 which can't see the lesion,

225 00:15:23.809 --> 00:15:27.269 and so they're trusting your algorithm

226 00:15:27.269 --> 00:15:34.190 to tell them exactly where to biopsy and you're also knowing that urologists are human and radiologists are human,

227 00:15:34.190 --> 00:15:38.549 and the outlines that they provide are not necessarily always completely accurate,

228 00:15:38.549 --> 00:15:42.190 and so you're dealing with a little bit of variability.

229 00:15:42.190 --> 00:15:44.740 But at the end of the day,

230 00:15:44.740 --> 00:15:52.370 the urologist puts that needle into the prostate into the part of the prostate that you told them to

231 00:15:52.370 --> 00:15:59.799 and then you go back and you look at the MRI to see whether or not they biopsied the right spot.

232 00:15:59.799 --> 00:16:00.190 Correct,

233 00:16:00.190 --> 00:16:05.620 you can do that. It's very interesting these cases that do have discordant results,

234 00:16:05.620 --> 00:16:12.990 which is expected we do go back and look at them and see what was missed in either case,

235 00:16:12.990 --> 00:16:14.929 but it's fascinating actually.

236 00:16:14.929 --> 00:16:21.139 If you look at the size of the gland compared to the size of the biopsy,

237 00:16:21.139 --> 00:16:23.860 it's something like .05% of the gland.

238 00:16:23.860 --> 00:16:30.220 That is all your sampling and many studies have actually shown that this targeting of biopsies

239 00:16:30.220 --> 00:16:36.179 is really the way to go because you get a much higher rate of detection of cancer that way.

240 00:16:36.179 --> 00:16:45.120 There's still a lot of variability in that and what's very interesting about the research that we've done here is we propose this novel fusion algorithm to hopefully

241 00:16:45.120 --> 00:16:52.269 map these lesions better and what we're able to do is here at Yale is we were able to see them in real time.

243 00:16:54.059 --> 00:16:58.230 Currently how they do it and then our method and we're able to

244 00:16:58.230 --> 00:17:00.929 see the variability in the targets itself.

245 00:17:00.929 --> 00:17:04.500 And then variability there, just the urologist looking at it,

246 00:17:04.500 --> 00:17:09.859 gave him some indication of how bad or incorrect that biopsy might be so while

247 00:17:09.859 --> 00:17:13.779 we weren't able to change a biopsy trajectory for the study,

248 00:17:13.779 --> 00:17:16.279 it gave an idea down the line

249 00:17:16.279 --> 00:17:19.849 of maybe this is why we missed this thing,

250 00:17:19.849 --> 00:17:26.279 because it was just a problem with sampling the wrong location because the wrong location

251 00:17:26.279 --> 00:17:29.470 was given.

252 00:17:29.470 --> 00:17:33.480 I'm sure that there are people who are listening to this going,

253 00:17:33.480 --> 00:17:38.819 I can't imagine that the wrong part of my prostate might be biopsied.

254 00:17:38.819 --> 00:17:48.839 How often is it inaccurate and how often is it inaccurate with the fusion technology versus how often is it inaccurate when you know the urologist goes in

255 00:17:48.839 --> 00:17:53.519 blind to do a biopsy under ultrasound of the thing that they can't see?

256 00:17:53.519 --> 00:17:56.519 I can't give you specific numbers on that.

257 00:17:56.519 --> 00:18:01.700 Studies do show that if you have a target presented by one of these devices

258 00:18:01.700 --> 00:18:07.099 you are much more likely to find that cancer that you were looking for,

259 00:18:07.099 --> 00:18:12.119 but again, that's something that's only available to a small number of institutions.

260 00:18:12.119 --> 00:18:20.609 Institutions that are larger are able to have these devices so traditionally a biopsy was taken in a just a regular systematic fashion.

261 00:18:20.609 --> 00:18:23.319 A urologist would only take 12 of them,

262 00:18:23.319 --> 00:18:28.720 and that's less of a game of chance.

263 00:18:28.720 --> 00:18:32.250 Like I said before, you're taking less than .05%

264 00:18:32.250 --> 00:18:37.349 of that prostate.

265 00:18:37.349 --> 00:18:42.450 You do end up with cases where you do find cancer where it wasn't suspected,

266 00:18:42.450 --> 00:18:45.990 certainly, and maybe that was just pure luck.

267 00:18:45.990 --> 00:18:48.950 But would you want to trust that, I don't know?

268 00:18:48.950 --> 00:18:53.089 You have much better chance of finding that cancer if you have these targets,

269 00:18:53.089 --> 00:18:55.460 even if these targets may not be 100%

270 00:18:55.460 --> 00:19:00.490 correct, it is much more likely that you're going to find it and be successful and have

271 00:19:00.490 --> 00:19:09.369 a better diagnosis.

272 00:19:09.369 --> 00:19:12.930 And so if on the MR are you see something suspicious and the radiologist says that's what we want to go after and you do the fusion algorithm and you target that thing and it comes back in,

273 00:19:12.930 --> 00:19:17.460 the pathologist says now it's benign. You talked before the break about

274 00:19:17.460 --> 00:19:19.940 discussing these cases in tumor board,

275 00:19:19.940 --> 00:19:25.599 tell us about what happens there and how you can get yourself either reassured that yeah,

276 00:19:25.599 --> 00:19:27.730 that really is benign, or

277 00:19:27.730 --> 00:19:30.910 we might have missed it even with our algorithm.

278 00:19:30.910 --> 00:19:33.039 That's what's great about the tumor

279 00:19:33.039 --> 00:19:38.349 board. It puts everybody that needs to make that decision in the room together.

280 00:19:38.349 --> 00:19:40.119 They're able to discuss it,

281 00:19:40.119 --> 00:19:46.130 so each specialist discusses what they see on either the imaging or the pathology,

282 00:19:46.130 --> 00:19:55.400 and then the urologist, what they saw during the procedure of the biopsy and it all kind of comes together to make one cohesive decision and

283 00:19:55.400 --> 00:19:58.410 a lot of time they come to some kind of consensus

284 00:19:58.410 --> 00:20:01.119 and the best plan is made for that patient.

285 00:20:01.119 --> 00:20:04.130 Often times if it is something that was not suspected,

286 00:20:04.130 --> 00:20:08.039 a patient will be placed on something that's called active surveillance.

287 00:20:08.039 --> 00:20:17.069 So they will be monitored more frequently for their care and the goal is that maybe if you missed it that first time by monitoring them actively,

288 00:20:17.069 --> 00:20:19.859 you'll be able to catch it a second time.

289 00:20:19.859 --> 00:20:21.450 Or if there's any progression.

290 00:20:21.450 --> 00:20:24.930 So if you missed it just by chance the first time,

291 00:20:24.930 --> 00:20:25.569 maybe they'll

292 00:20:25.569 --> 00:20:31.269 be more likely to see it the next time with all of the talk of AI,

293 00:20:31.269 --> 00:20:34.440 and there talk of AI in everything these days.

294 00:20:34.440 --> 00:20:36.660 I wonder about the downside of AI.

295 00:20:36.660 --> 00:20:39.200 I mean, certainly cost is likely an issue,

296 00:20:39.200 --> 00:20:41.109 and with health care costs rising

297 00:20:41.109 --> 00:20:51.220 I can't imagine that this is any cheaper or just as expensive as doing a regular biopsy, talk about the cost of the technology and the other downsides to AI.

298 00:20:51.220 --> 00:20:53.579 A we discussed before,

299 00:20:53.579 --> 00:20:58.630 AI, algorithms, or any kind of tools could be a real efficiency for clinicians.

300 00:20:58.630 --> 00:21:03.019 It could help them make decisions in an easier way, a cheaper way.

301 00:21:03.019 --> 00:21:09.079 The problem with training these algorithms is they are only as good as the data that you put in.

302 00:21:09.079 --> 00:21:11.099 There's the adage, garbage in,

303 00:21:11.099 --> 00:21:17.420 garbage out. So if you don't train these things with well annotated data or something that's really noisy,

304 00:21:17.420 --> 00:21:19.869 you're not going to get anything useful.

305 00:21:19.869 --> 00:21:26.170 That's a problem. Another inherent problem is this is they are potentially biased to whatever you trained on.

306 00:21:26.170 --> 00:21:31.420 So just for example, some of my own research I had 300 datasets from Yale,

307 00:21:31.420 --> 00:21:36.670 300 from Stanford. We trained an algorithm on one and ran it on the other.

308 00:21:36.670 --> 00:21:45.819 It didn't work. Shocking, we had perfect performance on the other site but something to realize is that these algorithms do not generalize well.

309 00:21:45.819 --> 00:21:51.059 You can't make a general inference as well as a human radiologist easily.

310 00:21:51.059 --> 00:21:55.660 Can a radiologist from Yale or Stanford easily tell what the prostate is?

311 00:21:55.660 --> 00:21:59.259 But this algorithm couldn't just because it was from a different

312 00:21:59.259 --> 00:22:01.819 location.

313 00:22:01.819 --> 00:22:03.910 So

314 00:22:03.910 --> 00:22:08.490 if I was living in California and I went to Stanford,

315 00:22:08.490 --> 00:22:13.500 and you did this fusion algorithm and did a biopsy,

316 00:22:13.500 --> 00:22:17.250 you'd be accurate. If I then went to Yale,

317 00:22:17.250 --> 00:22:18.920 you use the same

318 00:22:18.920 --> 00:22:21.000 algorithm, and it would be inaccurate?

319 00:22:21.970 --> 00:22:28.329 Potentially, yes.

320 00:22:28.329 --> 00:22:28.619 Then that means that you would have to retrain this algorithm for every new center that you plan on using it in, correct?

321 00:22:28.619 --> 00:22:30.349 That is an active area of research.

322 00:22:30.349 --> 00:22:34.109 Actually, people are looking at ways that they can either retrain things faster,

323 00:22:34.109 --> 00:22:37.579 or that they can just make these algorithms better from the start.

324 00:22:37.579 --> 00:22:42.779 Whether it's something you do to the data from the beginning of the pipeline and put it in,

325 00:22:42.779 --> 00:22:46.819 that can have a much better effect on your actual training of these things,

326 00:22:46.819 --> 00:22:50.869 but the problem you run into is what happens if somebody updates their software.

327 00:22:50.869 --> 00:22:54.049 You could just make your algorithm obsolete at that very moment,

328 00:22:54.049 --> 00:22:56.069 you have to retrain from scratch.

329 00:22:56.069 --> 00:23:01.950 So the most valuable thing again is what's the data that you're putting in here and how much of it.

330 00:23:01.950 --> 00:23:03.420 And that's really the key,

331 00:23:03.420 --> 00:23:05.740 and so are you able to

332 00:23:05.740 --> 00:23:11.700 use that data in a good way that can be applied throughout the entire population across all sites,

333 00:23:11.700 --> 00:23:13.349 in hospitals, in the US,

334 00:23:13.349 --> 00:23:14.349 in the world.

335 00:23:15.039 --> 00:23:21.569 Because one would think that if you are looking at an MR image.

336 00:23:21.569 --> 00:23:25.230 at Stanford you would be able to see what you see.

337 00:23:25.230 --> 00:23:31.940 You could take the same MR image and show it to a radiologist at Yale and they would see the same thing.

338 00:23:31.940 --> 00:23:35.299 It's like a photograph that I think a lot of this

339 00:23:35.299 --> 00:23:38.960 has to do with the misnomer of the name of artificial intelligence.

340 00:23:38.960 --> 00:23:47.500 Those of us who really work with the technology, we kind of cringe at that name because we know that there's no actual intelligence within the model itself.

341 00:23:47.500 --> 00:23:52.380 All the intelligence comes in from the data that people who created the data, that radiologists,

342 00:23:52.380 --> 00:23:54.819 the pathologist, the urologist, who created the data.

343 00:23:54.819 --> 00:23:56.380 That's where the intelligence is.

344 00:23:56.380 --> 00:23:58.250 So really it's just machine learning.

345 00:23:58.250 --> 00:24:01.680 This machine is learning to do something that a radiologist does,

346 00:24:01.680 --> 00:24:05.740 but it is not good at tasks that humans are really good at,

347 00:24:05.740 --> 00:24:07.299 which is making generalizable performance,

348 00:24:07.299 --> 00:24:11.359 making inferences very easily that apply to things that it has never seen.

349 00:24:11.359 --> 00:24:15.720 That's what the problem in our domain is called over training to the data.

350 00:24:15.720 --> 00:24:18.220 It's only good at things that I've seen,

351 00:24:18.220 --> 00:24:21.339 and it can't recognize something that has never seen before,

352 00:24:21.339 --> 00:24:23.519 which is a particular challenge when there's

353 00:24:23.519 --> 00:24:25.079 any kind of pathology, right?

355 00:24:27.119 --> 00:24:37.829 I'm just struggling with this because I think about the utility of the technology, before the break we said one of the utilities is really to help

356 00:24:37.829 --> 00:24:41.400 radiologists, who may not be specific to prostate cancer,

357 00:24:41.400 --> 00:24:45.690 who maybe the technology can help them to get better,

358 00:24:45.690 --> 00:24:56.400 but in that case you would be taking this technology out to a site that presumably didn't train it because it was trained by the experts at another

359 00:24:56.400 --> 00:25:00.339 site. But one would hope that it would be accurate at that second site,

360 00:25:00.339 --> 00:25:04.740 and if you train it at Stanford and tested at Yale or vice versa,

361 00:25:04.740 --> 00:25:06.390 and you didn't get any accuracy,

362 00:25:06.390 --> 00:25:09.410 I wonder what would happen if you trained at Yale,

363 00:25:09.410 --> 00:25:11.339 and then you took it out to,

364 00:25:11.339 --> 00:25:20.009 you know, Tuktoyaktuk, and for anybody who's wondering that's a small town in Canada, and it might not work.

365 00:25:20.009 --> 00:25:22.349 That's absolutely true. But fear not,

366 00:25:22.349 --> 00:25:22.740 that

367 00:25:22.740 --> 00:25:28.569 is something that the machine intelligence and machine learning people are trying to work on.

368 00:25:28.569 --> 00:25:33.630 I mean, that is probably the big problem right now in the community.

369 00:25:33.630 --> 00:25:36.740 This is especially true in the medical field.

370 00:25:36.740 --> 00:25:43.819 A lot of research that has gone on in this machine learning artificial intelligence has come out of stuff

371 00:25:43.819 --> 00:25:48.609 that Google and Apple and all these other big companies are doing with photographs,

372 00:25:48.609 --> 00:25:50.200 images, those are all good.

373 00:25:50.200 --> 00:25:57.220 They generalize fairly well. But what happens when human life is on the line when you're trying to work with these algorithms,

374 00:25:57.220 --> 00:26:02.000 there's a certain bar that we need to clear that is much higher than that.

375 00:26:02.000 --> 00:26:03.920 So

376 00:26:03.920 --> 00:26:07.430 we have to be very careful with what we're doing,

377 00:26:07.430 --> 00:26:13.849 and it is, again, it's a very active field of research that I think is probably the most critical thing.

378 00:26:13.849 --> 00:26:22.230 And it's also not to say that all these other companies that have their algorithms to recognize your cats and your dogs,

379 00:26:22.230 --> 00:26:25.660 they face the exact same problem with their cameras.

380 00:26:25.660 --> 00:26:29.089 What if they change their lens on their camera?

381 00:26:29.089 --> 00:26:31.380 Most likely that algorithm is going

382 00:26:31.380 --> 00:26:35.569 to have to be retrained to recognize your cat or dog.

383 00:26:35.569 --> 00:26:37.470 Interesting, what about the cost?

384 00:26:37.470 --> 00:26:38.230 I see

385 00:26:38.230 --> 00:26:42.039 that you sidestep that issue that I raised a while ago.

386 00:26:42.039 --> 00:26:44.650 It's actually the software.

387 00:26:44.650 --> 00:26:55.210 Hardware is relatively cheap. The innovations that came out the hardware are actually what really enabled this revolution that we're having now in this machine intelligence,

388 00:26:55.210 --> 00:26:58.049 it basically came out of video gaming.

389 00:26:58.049 --> 00:27:05.359 The graphics processing units of your computers are now able to crunch millions of calculations within a second,

390 00:27:05.359 --> 00:27:07.390 and that's what's really enabled

391 00:27:07.390 --> 00:27:11.849 this, and what's fascinating is a lot of people have called this

392 00:27:11.849 --> 00:27:15.210 the democratization of machine learning or machine intelligence

393 00:27:15.210 --> 00:27:23.069 because Google and Facebook have made these algorithms in these toolkits available that high school students can take.

394 00:27:23.069 --> 00:27:25.750 They can build these deep learning models.

395 00:27:25.750 --> 00:27:33.009 These artificial neural networks and get solutions to problems that we previously had to engineer these complex models with.

396 00:27:33.009 --> 00:27:43.180 And now you can just take these tools out of the box and you can run it and they can get an answer that's surprisingly good.

397 00:27:43.180 --> 00:27:47.759 But what's really lacking is the understanding of what that model can do,

398 00:27:47.759 --> 00:27:53.390 and also what are some other things that we can do as researchers or as clinicians?

399 00:27:53.390 --> 00:27:59.720 What can we add that we already know to improve these models in the training of these things?

400 00:27:59.720 --> 00:28:06.059 And so that's the challenge, bringing in things that can help them learn in a better way.

401 00:28:06.059 --> 00:28:06.759 And so

402 00:28:06.759 --> 00:28:08.880 where are we on that front?

403 00:28:09.829 --> 00:28:15.170 Well, we are in the midst of it, there's a big investment in this.

404 00:28:15.170 --> 00:28:18.509 Lot of companies are investing in this and it's just

405 00:28:18.509 --> 00:28:22.190 burgeoning right now where there's very rapid uptake and research.

406 00:28:22.190 --> 00:28:25.529 Everybody is doing it now everybody's jumping on the bandwagon.

407 00:28:25.529 --> 00:28:33.880 There's tons of money out there and I think we're at the point where now we really need to evaluate how good these models are.

408 00:28:33.880 --> 00:28:37.220 The evaluation of the validation is going to be critical.

409 00:28:37.220 --> 00:28:41.230 There's a lot of hype right now and trying to apply this,

410 00:28:41.230 --> 00:28:45.920 especially to medicine, but I think we need to be very careful on how we apply this.

411 00:28:45.920 --> 00:28:47.750 And there's also the questions of

412 00:28:47.750 --> 00:28:50.369 is there bias? Are the ethics issues involved in this.

413 00:28:50.369 --> 00:28:51.940 Where does the data come from?

414 00:28:51.940 --> 00:28:53.250 How important is that data?

415 00:28:53.250 --> 00:28:56.400 Again, there's a lot of questions that need to be answered now,

416 00:28:56.400 --> 00:28:58.759 and it's a very exciting time in the field.

417 00:28:59.339 --> 00:29:07.240 Doctor John Onofrey is assistant professor of radiology and biomedical imaging and of urology at Yale School of Medicine.

418 00:29:07.240 --> 00:29:15.880 If you have questions, the address is canceranswers@yale.edu and past editions of the program are available in audio and written form at Yalecancercenter.org.

419 00:29:15.880 --> 00:29:24.048 We hope you'll join us next week to learn more about the fight against cancer here on Connecticut Public Radio.