WEBVTT

1 00:00:00.050 --> 00:00:03.230 - [Fan] David Benkeser who is an assistant professor

2 00:00:03.230 --> 00:00:06.530 at the department of biostatistics and bioinformatics

3 00:00:06.530 --> 00:00:08.700 at Emory University.

4 00:00:08.700 --> 00:00:11.210 Dr. Benkeser got his PhD in biostatistics

5 00:00:11.210 --> 00:00:12.700 from University of Washington

6 00:00:12.700 --> 00:00:14.720 and had his post-doctoral fellowship

7 00:00:14.720 --> 00:00:17.423 from University of California at Berkeley.

8 00:00:18.270 --> 00:00:21.660 Dr. Benkeser is an expert in methods for machine learning

9 00:00:21.660 --> 00:00:24.070 and non-parametric statistical inference.

10 00:00:24.070 --> 00:00:25.760 He has made important contributions

11 00:00:25.760 --> 00:00:27.900 to integrate machine learning methods

12 00:00:27.900 --> 00:00:31.140 to draw causal inferences with observational data.

13 00:00:31.140 --> 00:00:33.660 He also has interesting work on preventative vaccines

14 00:00:33.660 --> 00:00:37.440 and HIV prevention, which he's going to share with us today.

15 00:00:37.440 --> 00:00:39.223 Welcome David, the floor is yours.

16 00:00:43.880 --> 00:00:45.810 - [David] Thanks, yeah, it's a great pleasure

17 00:00:45.810 --> 00:00:47.520 to be here today.

18 00:00:47.520 --> 00:00:51.280 Well, here today, but with you guys today giving this talk.

19 00:00:51.280 --> 00:00:54.590 So I did see that I think Tony Fauci

20 00:00:54.590 --> 00:00:57.620 spoke at Yale yesterday, so it was very nice of you Fan

21 00:00:57.620 --> 00:01:00.340 to book Tony Fauci as my opening act

22 00:01:00.340 --> 00:01:04.270 and I'll try not to disappoint him with my followup.

23 00:01:04.270 --> 00:01:07.120 So the talk I'm giving today is a very high-level talk.

24 00:01:07.120 --> 00:01:10.070 So the title is statistics and COVID-19 vaccine development,

25 00:01:10.070 --> 00:01:12.280 but it's really a talk mostly

26 00:01:12.280 --> 00:01:14.700 about COVID-19 vaccine development.

27 00:01:14.700 --> 00:01:18.710 There's not math until maybe slide like 29 out of 30.

28 00:01:18.710 --> 00:01:19.960 So really these are sort of

29 00:01:19.960 --> 00:01:23.010 just the high-level issues that have come up

30 00:01:23.010 --> 00:01:28.010 as I've worked with companies and government organizations

31 00:01:28.020 --> 00:01:30.040 on COVID-19 vaccine development.

32 00:01:30.040 --> 00:01:31.790 So I think there's a lot of really interesting stuff

33 00:01:31.790 --> 00:01:35.370 here and really, really glad to share it with you today.

34 00:01:35.370 --> 00:01:38.679 So if you want to kind of slide along with

35 00:01:38.679 --> 00:01:41.140 the slides they're available on GitHub

36 00:01:41.140 --> 00:01:43.190 so there's a link at the bottom there,

37 00:01:43.190 --> 00:01:44.267 and you can click on that

38 00:01:44.267 --> 00:01:45.870 and that'll pull up the HTML slide back,

39 00:01:45.870 --> 00:01:48.290 and I have sort of references hyperlinked in there.

40 00:01:48.290 --> 00:01:49.580 So that's an easy way to access

41 00:01:49.580 --> 00:01:51.890 the references there as well.

42 00:01:51.890 --> 00:01:54.520 Okay so I'm going to start just kind of talking

43 00:01:54.520 --> 00:01:57.750 about the biology a little bit of SARS-CoV-2,

44 00:01:57.750 --> 00:02:01.110 and segue into sort of how we can think about

45 00:02:01.110 --> 00:02:02.790 developing vaccines that will prevent

46 00:02:02.790 --> 00:02:05.550 an infection and COVID-19 disease.

47 00:02:05.550 --> 00:02:09.330 And so this is a nice little graphic that I ripped off

48 00:02:09.330 --> 00:02:11.740 from The Washington Post, who's very much better

49 00:02:11.740 --> 00:02:13.290 at making these cutesy little graphics

50 00:02:13.290 --> 00:02:16.140 than I am using PowerPoint or something.

51 00:02:16.140 --> 00:02:17.450 So let's kind of walk through this.

52 00:02:17.450 --> 00:02:19.520 And the goal here is to try to understand,

53 00:02:19.520 --> 00:02:21.620 you know, how SARS-CoV-2 is infecting your cells,

54 00:02:21.620 --> 00:02:22.920 how it's replicating,

55 00:02:22.920 --> 00:02:25.040 and then to understand what the mechanisms

56 00:02:25.040 --> 00:02:27.570 that immunological mechanisms of the vaccine are

57 00:02:27.570 --> 00:02:29.930 that can potentially block that infection

58 00:02:29.930 --> 00:02:31.120 and prevent clinical disease.

59 00:02:31.120 --> 00:02:32.950 So we'll just go quickly through this

60 00:02:32.950 --> 00:02:36.320 and this is sort of the story for most viruses, right?

61 00:02:36.320 --> 00:02:39.040 Is that viruses are really just genetic material

62 00:02:39.040 --> 00:02:42.162 in this case RNA that's wrapped up in the glycoprotein.

63 00:02:42.162 --> 00:02:45.190 So it's genetic material wrapped up in a protein.

64 00:02:45.190 --> 00:02:48.560 And so for SARS-CoV-2 you may have heard of a couple

65 00:02:48.560 --> 00:02:50.900 of these proteins in particular, the spike protein will play

66 00:02:50.900 --> 00:02:53.680 a large role when we talk about a vaccine development

67 00:02:53.680 --> 00:02:56.100 and why is this spike protein so important?

68 00:02:56.100 --> 00:02:58.170 Well, that's the guy that sort of latches

69 00:02:58.170 --> 00:03:01.375 onto your cell and it does that through this ACE2 pathway

70 00:03:01.375 --> 00:03:05.240 and it grabs onto your cell and insert itself inside

71 00:03:05.240 --> 00:03:07.110 you cell and once it's inside

72 00:03:07.110 --> 00:03:08.860 it releases its genetic material, right?

73 00:03:08.860 --> 00:03:11.410 It releases its RNA and kind of tricks

74 00:03:11.410 --> 00:03:13.430 your cell into replicating the virus, right?

75 00:03:13.430 --> 00:03:17.060 So that your cell is producing new copies of this virus,

76 00:03:17.060 --> 00:03:18.970 they're pieced together out of proteins that are released

77 00:03:18.970 --> 00:03:21.570 into your bloodstream to go infect more cells

78 00:03:21.570 --> 00:03:22.980 and more people.

79 00:03:22.980 --> 00:03:24.810 Okay so that's sort of the infection process

80 00:03:24.810 --> 00:03:26.740 and where along the lines do you know

81 00:03:26.740 --> 00:03:28.800 vaccines sort of halt this?

82 00:03:28.800 --> 00:03:30.670 So I'll walk through a few different

83 00:03:30.670 --> 00:03:33.130 of the major vaccine constructs that are being used

84 00:03:33.130 --> 00:03:34.860 for SARS-CoV-2 vaccines,

85 00:03:34.860 --> 00:03:37.080 and the details aren't super important here,

86 00:03:37.080 --> 00:03:38.320 but I do think it's sort of helpful

87 00:03:38.320 --> 00:03:39.953 to have a high level overview in comparison, right?

88 00:03:39.953 --> 00:03:42.970 Because there's so many vaccine products being developed,

89 00:03:42.970 --> 00:03:45.180 at least having some point of biological comparison

90 00:03:45.180 --> 00:03:47.270 of how they're working is useful.

91 00:03:47.270 --> 00:03:48.430 So to walk through these slides,

92 00:03:48.430 --> 00:03:50.440 all of these slides are basically going to be the same

93 00:03:50.440 --> 00:03:52.370 on the right hand side of the slide

94 00:03:52.370 --> 00:03:53.960 and how they're gonna differ is what goes

95 00:03:53.960 --> 00:03:55.580 into the vaccine on the left-hand side.

96 00:03:55.580 --> 00:03:57.837 So let's actually start on the right-hand side, right?

97 00:03:57.837 --> 00:04:00.250 And talk a little bit about immunology, right?

98 00:04:00.250 --> 00:04:02.997 And how your body tries to fight off infection.

99 00:04:02.997 --> 00:04:04.740 And we have a couple of different mechanisms

100 00:04:04.740 --> 00:04:06.020 of your immune system to do that.

101 00:04:06.020 --> 00:04:09.500 So there's a kind of T-cell responses, cytotoxic T-cells.

102 00:04:09.500 --> 00:04:12.210 So those are T-cells that recognize cells in your body

103 00:04:12.210 --> 00:04:13.850 that have been infected with a pathogen

104 00:04:13.850 --> 00:04:15.210 and destroy those cells, right?

105 00:04:15.210 --> 00:04:17.110 Because the cells are producing copies of the virus,

106 00:04:17.110 --> 00:04:18.840 releasing in the bloodstream.

107 00:04:18.840 --> 00:04:21.400 So if we're able to destroy infected cells,

108 00:04:21.400 --> 00:04:24.290 we can potentially stop infection, prevent disease,

109 00:04:24.290 --> 00:04:25.580 and then another key response

110 00:04:25.580 --> 00:04:27.510 that your immune system has is through antibodies.

111 00:04:27.510 --> 00:04:29.560 And that's sort of what's on the bottom here

112 00:04:29.560 --> 00:04:33.860 and is that B cells are able to produce antibodies.

113 00:04:33.860 --> 00:04:36.100 And what those antibodies do is they basically grab

114 00:04:36.100 --> 00:04:38.280 onto these surface proteins, right?

115 00:04:38.280 --> 00:04:40.280 So remember we talked about the spike protein,

116 00:04:40.280 --> 00:04:42.730 and what antibodies do is basically just bind onto that

117 00:04:42.730 --> 00:04:45.520 and sit there and so neutralizing antibodies.

118 00:04:45.520 --> 00:04:47.480 So there's two classes of antibodies that are kind

119 00:04:47.480 --> 00:04:48.470 of relevant for vaccines.

120 00:04:48.470 --> 00:04:50.113 So neutralizing antibodies really, you're just doing that.

121 00:04:50.113 --> 00:04:52.787 They're gonna sit on all of those spike proteins

122 00:04:52.787 --> 00:04:55.130 and because they're sitting there now the virus can't grab

123 00:04:55.130 --> 00:04:57.010 onto your cells to infect them.

124 00:04:57.010 --> 00:05:00.680 There's also binding antibodies, which are somewhat

125 00:05:00.680 --> 00:05:02.510 considered to be less important in this context,

126 00:05:02.510 --> 00:05:04.830 but what those guys do is bind onto those surface proteins,

127 00:05:04.830 --> 00:05:06.840 they don't neutralize the virus itself,

128 00:05:06.840 --> 00:05:09.160 but they send out chemical signals to other cells

129 00:05:09.160 --> 00:05:11.025 in your body that say, hey, here's a virus.

130 00:05:11.025 --> 00:05:13.030 Please come eat it for me.

131 00:05:13.030 --> 00:05:15.280 So those are the sort of antibody classes response

132 00:05:15.280 --> 00:05:16.113 that you can have.

133 00:05:16.113 --> 00:05:18.240 So there's these two sort of immune mechanisms

134 00:05:18.240 --> 00:05:22.460 that we have to neutralize infections by viruses.

135 00:05:22.460 --> 00:05:24.720 How do they learn to neutralize them?

136 00:05:24.720 --> 00:05:25.980 Well, there's this sort of middleman.

137 00:05:25.980 --> 00:05:27.790 So we're moving just to this middle panel here

138 00:05:27.790 --> 00:05:29.630 with these APC cells,

139 00:05:29.630 --> 00:05:31.830 so these antigen presenting cells, right?

140 00:05:31.830 --> 00:05:33.240 Those are the guys that what they're doing

141 00:05:33.240 --> 00:05:35.890 is basically digesting little bits

142 00:05:35.890 --> 00:05:40.347 of the virus in this case of the surface protein, right?

143 00:05:40.347 --> 00:05:42.540 And they're teaching or training your immune system

144 00:05:42.540 --> 00:05:44.250 to recognize that pathogen, right?

145 00:05:44.250 --> 00:05:46.790 So they're the ones that go and talk to the T cells,

146 00:05:46.790 --> 00:05:48.120 talk to the B cells and say,

147 00:05:48.120 --> 00:05:50.580 here's that how this virus looks,

148 00:05:50.580 --> 00:05:53.220 please go produce some antibodies or please recognize cells

149 00:05:53.220 --> 00:05:54.680 that have been infected with this

150 00:05:54.680 --> 00:05:56.330 and neutralize them for me.

151 00:05:56.330 --> 00:05:58.990 So really again, the whole right side of this plot

152 00:05:58.990 --> 00:06:00.220 is about your immune system.

153 00:06:00.220 --> 00:06:02.960 This is the way your immune system fights off infection.

154 00:06:02.960 --> 00:06:04.780 And what's different between this slide

155 00:06:04.780 --> 00:06:07.820 and the next few slides is basically how we present

156 00:06:07.820 --> 00:06:09.280 pieces of the pathogen pieces

157 00:06:09.280 --> 00:06:11.720 of the virus to these APCs, right?

158 00:06:11.720 --> 00:06:14.440 So how do we get these APCs, the material that they need

159 00:06:14.440 --> 00:06:18.860 for you to mount an immune response against SARS-CoV-2?

160 00:06:18.860 --> 00:06:20.700 And so the first class of vaccines

161 00:06:20.700 --> 00:06:22.830 I'll describe are nucleic acid vaccines.

162 00:06:22.830 --> 00:06:25.040 And so I'm talking about first

163 00:06:25.040 --> 00:06:27.050 because they're sort of the first wave of vaccines

164 00:06:27.050 --> 00:06:29.130 that are in phase three trials in the US.

165 00:06:29.130 --> 00:06:32.847 So Moderna and Pfizer, who are probably the most advanced

166 00:06:32.847 --> 00:06:36.897 candidates for US licensure are both mRNA vaccines.

167 00:06:36.897 --> 00:06:38.630 And so how are those vaccines made?

168 00:06:38.630 --> 00:06:41.830 Well, we take a little bit of messenger RNA,

169 00:06:41.830 --> 00:06:43.820 a little bit of viral genetic material,

170 00:06:43.820 --> 00:06:45.270 and wrap that in a lipid shell, right?

171 00:06:45.270 --> 00:06:46.710 That's the construct of the vaccine.

172 00:06:46.710 --> 00:06:49.190 And when you're injected that lipid shell

173 00:06:49.190 --> 00:06:51.240 latches onto your cell, right?

174 00:06:51.240 --> 00:06:53.710 Delivers that mRNA into your cell,

175 00:06:53.710 --> 00:06:55.490 just like a natural infection, right?

176 00:06:55.490 --> 00:06:57.470 Remember the SARS-CoV-2 grabbed onto your cell

177 00:06:57.470 --> 00:07:00.440 and inserted itself and then made copies of itself.

178 00:07:00.440 --> 00:07:02.990 So what is the mRNA doing once it's in your cell,

179 00:07:02.990 --> 00:07:04.580 it's actually just making copies

180 00:07:04.580 --> 00:07:06.810 of the spike protein itself, right?

181 00:07:06.810 --> 00:07:11.040 So you're manufacturing this protein within your own cells

182 00:07:11.040 --> 00:07:14.370 that are then released for these APCs to detect.

183 00:07:14.370 --> 00:07:16.890 So this is how we're getting these APCs,

184 00:07:16.890 --> 00:07:18.940 spike protein with an mRNA vaccine.

185 00:07:18.940 --> 00:07:21.790 We're basically using your cells as a warehouse

186 00:07:21.790 --> 00:07:23.637 to produce the antigen of the vaccine

187 00:07:23.637 --> 00:07:27.670 and so this is a really cool idea and a new idea, right?

188 00:07:27.670 --> 00:07:31.010 So, am mRNA or DNA vaccine has never been licensed before

189 00:07:31.010 --> 00:07:34.170 and that's not to say that we tried many times and failed.

190 00:07:34.170 --> 00:07:36.240 It's just to say that this is a very new technology,

191 00:07:36.240 --> 00:07:39.520 and it's sort of interesting that it's kind of come

192 00:07:39.520 --> 00:07:41.480 to the forefront in this context.

193 00:07:41.480 --> 00:07:43.940 So why do we like mRNA vaccines?

194 00:07:43.940 --> 00:07:45.610 Well, they're very fast to manufacturer.

195 00:07:45.610 --> 00:07:47.710 We'll talk about some of the other vaccine constructs

196 00:07:47.710 --> 00:07:50.210 where we're making this spike protein in a lab,

197 00:07:50.210 --> 00:07:52.250 and that is a long and arduous.

198 00:07:52.250 --> 00:07:53.950 It needs to be very careful process

199 00:07:53.950 --> 00:07:55.290 and when we're thinking about scaling up

200 00:07:55.290 --> 00:07:59.130 vaccine manufacturing, mRNA vaccines are very appealing

201 00:07:59.130 --> 00:08:01.770 in that sense, you can manufacture them

202 00:08:01.770 --> 00:08:03.300 very quickly at scale.

203 00:08:03.300 --> 00:08:04.870 They don't require a cold chain

204 00:08:04.870 --> 00:08:09.520 and so that's another great advantage these vaccines enjoy

205 00:08:09.520 --> 00:08:11.710 in terms of thinking about vaccine deployment,

206 00:08:11.710 --> 00:08:15.320 particularly in developing world settings.

207 00:08:15.320 --> 00:08:17.000 But again, this is a brand new technology.

208 00:08:17.000 --> 00:08:18.260 We don't have any safety data

209 00:08:18.260 --> 00:08:20.410 from past vaccines with this construct.

210 00:08:20.410 --> 00:08:22.400 We don't have any efficacy data.

211 00:08:22.400 --> 00:08:24.700 So, it's sort of an open question in the field

212 00:08:24.700 --> 00:08:26.850 as to how well these things are gonna work.

213 00:08:27.710 --> 00:08:30.117 So moving to sort of more classical, constructive vaccines

214 00:08:30.117 --> 00:08:32.680 and viral vector vaccines.

215 00:08:32.680 --> 00:08:34.150 So again, the right side of this picture

216 00:08:34.150 --> 00:08:35.070 is exactly the same.

217 00:08:35.070 --> 00:08:38.110 The story is how do we get an APC the right antigen?

218 00:08:38.110 --> 00:08:41.390 How do we show an APC a little bit of the spike protein?

219 00:08:41.390 --> 00:08:45.482 So a viral vector vaccine, right?

220 00:08:45.482 --> 00:08:48.960 Is going to take a different virus and splice

221 00:08:48.960 --> 00:08:52.130 a little bit of SARS-CoV-2 into that virus, okay.

222 00:08:52.130 --> 00:08:55.700 So for example, AstraZeneca, that's the Oxford that you may

223 00:08:55.700 --> 00:08:58.470 have heard of, they take a chimpanzee adenovirus,

224 00:08:58.470 --> 00:09:02.210 that's like, it's a virus that causes the common cold

225 00:09:02.210 --> 00:09:04.250 in chimpanzees and they splice in a little bit

226 00:09:04.250 --> 00:09:09.230 of SARS-CoV-2 into that and so that sort of host virus,

227 00:09:09.230 --> 00:09:12.120 that adenovirus holds genetic material

228 00:09:12.120 --> 00:09:15.690 infects your cells and your cells then produce the antigen.

229 00:09:15.690 --> 00:09:19.790 They produce the spike protein of SARS-CoV-2.

230 00:09:19.790 --> 00:09:22.420 So AstraZeneca and Janssen are using this construct again,

231 00:09:22.420 --> 00:09:26.350 both with adenoviruses, a very common virus vector.

232 00:09:26.350 --> 00:09:29.100 And again, we like these types of vaccines

233 00:09:29.100 --> 00:09:31.720 because they're quick to manufacturer,

234 00:09:31.720 --> 00:09:34.130 but a challenge of them is that your body

235 00:09:34.130 --> 00:09:36.740 can sort of develop separate immune responses

236 00:09:36.740 --> 00:09:39.450 against the vector itself, right?

237 00:09:39.450 --> 00:09:41.930 So you can develop a separate immune response

238 00:09:41.930 --> 00:09:44.510 against say an adenovirus right?

239 00:09:44.510 --> 00:09:47.360 Such that your body neutralizes those adenoviruses

240 00:09:47.360 --> 00:09:49.910 before they're able to infect your cells

241 00:09:49.910 --> 00:09:51.720 and produce the SARS-CoV-2 antigen.

242 00:09:51.720 --> 00:09:55.074 So we do see tendency a kind of faster waning

243 00:09:55.074 --> 00:09:58.470 vaccine effects with this class of vaccines.

244 00:09:58.470 --> 00:10:00.470 So moving on to subunit vaccine.

245 00:10:00.470 --> 00:10:02.940 So this is NovaVax and Sanofi's vaccine

246 00:10:02.940 --> 00:10:04.800 will be subunit vaccines

247 00:10:04.800 --> 00:10:06.250 and this is where I kind of mentioned before

248 00:10:06.250 --> 00:10:09.790 actually what happens here is these spike proteins

249 00:10:09.790 --> 00:10:11.700 or whatever the antigen is,

250 00:10:11.700 --> 00:10:14.380 is created and purified in a lab.

251 00:10:14.380 --> 00:10:17.010 So they actually use insect cells

252 00:10:17.010 --> 00:10:19.900 that they infect with SARS-CoV-2,

253 00:10:19.900 --> 00:10:23.310 those insect cells then produce the antigen that's purified

254 00:10:23.310 --> 00:10:25.920 and that's what goes into the vaccine

255 00:10:25.920 --> 00:10:27.470 are those protein subunits, right?

256 00:10:27.470 --> 00:10:29.800 So there we're just directly giving you the spike protein

257 00:10:29.800 --> 00:10:33.290 that we've grown outside of the host

258 00:10:33.290 --> 00:10:36.950 and that's how we're getting these APCs, those antigens.

259 00:10:36.950 --> 00:10:40.080 And so this is a commonly used vaccine construct.

260 00:10:40.080 --> 00:10:42.160 So the hep B vaccine is highly effective.

261 00:10:42.160 --> 00:10:43.890 HPV vaccine is highly effective.

262 00:10:43.890 --> 00:10:45.987 That's the construct of these, but the downside of course

263 00:10:45.987 --> 00:10:50.580 to it, so it's a well-trodden way of developing vaccines.

264 00:10:50.580 --> 00:10:52.750 But the downside is that they're slower to manufacturer.

265 00:10:52.750 --> 00:10:55.180 There's this whole process where we have to cultivate

266 00:10:55.180 --> 00:11:00.180 and grow these viruses in a lab, we have to purify them,

267 00:11:00.239 --> 00:11:03.930 and moreover they often also require an adjuvant.

268 00:11:03.930 --> 00:11:06.640 So that's really just sort of adding something a little bit

269 00:11:06.640 --> 00:11:10.660 extra that stimulates a better immune response in your body.

270 00:11:10.660 --> 00:11:12.760 So basically at the site of injection,

271 00:11:12.760 --> 00:11:13.930 it's something that increases

272 00:11:13.930 --> 00:11:15.770 your inflammatory response actually

273 00:11:15.770 --> 00:11:17.900 to kind of stimulate your immune system

274 00:11:17.900 --> 00:11:19.720 into recognizing those antigens

275 00:11:19.720 --> 00:11:22.360 and developing an immune response against them.

276 00:11:22.360 --> 00:11:24.360 So there's subunit vaccines.

277 00:11:24.360 --> 00:11:27.447 So the fourth class here is a weakened/inactivated vaccine.

278 00:11:27.447 --> 00:11:30.240 And so this is, I think, what most people like what

279 00:11:30.240 --> 00:11:32.980 my grandparents probably think all vaccines are,

280 00:11:32.980 --> 00:11:35.270 is basically we take a pathogen

281 00:11:35.270 --> 00:11:38.530 and we weaken it in some way, or we kill it, right?

282 00:11:38.530 --> 00:11:40.180 And then that's the construct of the vaccine

283 00:11:40.180 --> 00:11:42.130 and that's what's injected into you.

284 00:11:42.130 --> 00:11:45.440 And we go through this similar process there

285 00:11:45.440 --> 00:11:47.280 that literally mimics natural infection, right?

286 00:11:47.280 --> 00:11:51.230 Where your cells are infected by this weakened form

287 00:11:51.230 --> 00:11:53.070 of the virus, the virus replicates,

288 00:11:53.070 --> 00:11:55.810 and that's how we get antigens to the APCs.

289 00:11:55.810 --> 00:11:57.540 So this is the construct used in of course

290 00:11:57.540 --> 00:12:01.400 some classic vaccines like MMR, polio vaccine,

291 00:12:01.400 --> 00:12:02.970 but again, it's slower manufacturing, right?

292 00:12:02.970 --> 00:12:04.490 Because we have to cultivate the virus

293 00:12:04.490 --> 00:12:07.280 in the lab and then it also requires adjuvants.

294 00:12:07.280 --> 00:12:10.110 So I don't think there's currently any plans

295 00:12:10.110 --> 00:12:12.150 to have US phase three trials

296 00:12:12.150 --> 00:12:15.480 of weaken inactivated vaccines, but there are in China.

297 00:12:15.480 --> 00:12:17.383 So Sinopharm and Sinovac vaccines

298 00:12:17.383 --> 00:12:19.033 were using this construct.

299 00:12:20.554 --> 00:12:23.630 So that's just a bit of a background in immunology

300 00:12:23.630 --> 00:12:26.460 and how all this works and how we think about preventing

301 00:12:26.460 --> 00:12:29.460 infection with SARS-CoV-2 and hopefully preventing

302 00:12:29.460 --> 00:12:32.090 clinical disease COVID-19 disease.

303 00:12:32.090 --> 00:12:33.740 So now we're gonna segue to talk a little bit

304 00:12:33.740 --> 00:12:35.590 about the vaccine development process, right?

305 00:12:35.590 --> 00:12:37.510 'Cause this has all happened extremely fast.

306 00:12:37.510 --> 00:12:40.340 So let's talk about sort of the process whereby

307 00:12:40.340 --> 00:12:43.310 vaccine products are typically brought to market, right.

308 00:12:43.310 --> 00:12:44.810 And what looks a little bit different

309 00:12:44.810 --> 00:12:49.080 about the COVID-19 vaccine development process?

310 00:12:49.080 --> 00:12:52.350 So this is a figure from a nice New England journal paper

311 00:12:52.350 --> 00:12:53.650 that's referenced at the bottom

312 00:12:53.650 --> 00:12:55.560 that's just talking about sort of what's different

313 00:12:55.560 --> 00:12:57.730 this go around in terms of how are we accelerating

314 00:12:57.730 --> 00:12:59.730 the vaccine development process.

315 00:12:59.730 --> 00:13:01.120 And so I think as biostatisticians,

316 00:13:01.120 --> 00:13:04.101 anyone who works on clinical trials is fairly familiar

317 00:13:04.101 --> 00:13:05.890 with the traditional paradigm

318 00:13:05.890 --> 00:13:08.350 for bringing products to market, right.

319 00:13:08.350 --> 00:13:11.020 It involves sort of a lot of R&D

320 00:13:11.020 --> 00:13:12.900 in the lab, preclinical work

321 00:13:12.900 --> 00:13:15.820 and then you start doing human trials in phase one,

322 00:13:15.820 --> 00:13:19.380 these are small dose finding safety trials,

323 00:13:19.380 --> 00:13:20.780 checking whether these vaccines

324 00:13:20.780 --> 00:13:22.940 generate any immune response.

325 00:13:22.940 --> 00:13:25.340 And then what we'll often do is in vaccine trials

326 00:13:25.340 --> 00:13:27.100 is run a small randomized trial.

327 00:13:27.100 --> 00:13:28.580 That's a phase two trial, right?

328 00:13:28.580 --> 00:13:31.060 We're we'll have a placebo control,

329 00:13:31.060 --> 00:13:33.855 maybe pick out a particularly high risk population

330 00:13:33.855 --> 00:13:35.280 and start to see if we're getting

331 00:13:35.280 --> 00:13:37.100 any efficacy signal, right?

332 00:13:37.100 --> 00:13:38.970 And this is a very deliberate process, right?

333 00:13:38.970 --> 00:13:41.000 Phase one typically advances very slowly.

334 00:13:41.000 --> 00:13:42.380 We have lots of safety concerns.

335 00:13:42.380 --> 00:13:45.420 Phase two, we think very hard about whether the efficacy

336 00:13:45.420 --> 00:13:47.750 signal was really worth it to advance a candidate to

337 00:13:47.750 --> 00:13:50.800 phase three and it's a very deliberate process, right?

338 00:13:50.800 --> 00:13:53.010 To get to this phase three licensure trial, right?

339 00:13:53.010 --> 00:13:55.560 So the phase three trial is the big one involving

340 00:13:55.560 --> 00:13:56.530 the most participants.

341 00:13:56.530 --> 00:13:59.900 It's a randomized controlled trial, right?

342 00:13:59.900 --> 00:14:02.060 Enrolling many, many subjects that's well powered

343 00:14:02.060 --> 00:14:04.460 to detect ethicacy signals and based on the results

344 00:14:04.460 --> 00:14:06.900 of that phase three trial and safety data

345 00:14:06.900 --> 00:14:08.260 that's been accumulated throughout

346 00:14:08.260 --> 00:14:09.830 this whole process, right.

347 00:14:09.830 --> 00:14:13.690 We're able to provide licensure ideally for a product.

348 00:14:13.690 --> 00:14:16.510 And so that's sort of the clinical development process,

349 00:14:16.510 --> 00:14:18.760 but also in the context of COVID vaccines

350 00:14:18.760 --> 00:14:19.640 it's important to think about

351 00:14:19.640 --> 00:14:21.377 the manufacturing process, right.

352 00:14:21.377 --> 00:14:23.050 And how that looks a little bit different.

353 00:14:23.050 --> 00:14:27.010 So typically right, companies are very sort of hesitant

354 00:14:27.010 --> 00:14:30.640 to scale up manufacturing before they know that they have

355 00:14:30.640 --> 00:14:31.980 a product that will be licensed, right.

356 00:14:31.980 --> 00:14:34.040 Which makes sense, you know, they're sort of risk averse.

357 00:14:34.040 --> 00:14:35.710 We don't want to start manufacturing a product

358 00:14:35.710 --> 00:14:38.780 that may ultimately be shot down by the FDA.

359 00:14:38.780 --> 00:14:40.320 So really large scale manufacturing

360 00:14:40.320 --> 00:14:43.670 is not happening until after product licensure.

361 00:14:43.670 --> 00:14:45.010 So what's happening with COVID vaccine

362 00:14:45.010 --> 00:14:48.620 is basically this whole long deliberate timeline

363 00:14:48.620 --> 00:14:51.650 is being compressed into a shorter time period.

364 00:14:51.650 --> 00:14:53.370 And so how do we do that?

365 00:14:53.370 --> 00:14:55.680 Well, basically what happens is we've collapsed

366 00:14:55.680 --> 00:14:57.650 the phase one and phase two trials, right?

367 00:14:57.650 --> 00:14:59.790 So we're doing small safety studies.

368 00:14:59.790 --> 00:15:01.320 We're checking whether these vaccines

369 00:15:01.320 --> 00:15:02.840 are generating immune responses,

370 00:15:02.840 --> 00:15:05.956 but we're really not doing that smaller efficacy study

371 00:15:05.956 --> 00:15:10.150 that is typical of vaccine development.

372 00:15:10.150 --> 00:15:12.710 And so we're collapsing the phase one and two process,

373 00:15:12.710 --> 00:15:14.770 the phase three process is where we're at, right.

374 00:15:14.770 --> 00:15:16.420 We're doing these large scale trials, right?

375 00:15:16.420 --> 00:15:18.607 Because we need robust efficacy data

376 00:15:18.607 --> 00:15:21.520 and we need robust safety data to gain licensure,

377 00:15:21.520 --> 00:15:23.870 but a big thing that has changed, so the clinical process

378 00:15:23.870 --> 00:15:26.770 yeah a little bit compressed, but mostly the same,

379 00:15:26.770 --> 00:15:27.740 the big thing that's changed

380 00:15:27.740 --> 00:15:29.820 is the manufacturing process, right.

381 00:15:29.820 --> 00:15:33.250 Is we wanna make sure that once a vaccine is licensed

382 00:15:33.250 --> 00:15:36.500 and is proven to be safe and effective that we're able

383 00:15:36.500 --> 00:15:38.530 to start distributing that vaccine immediately.

384 00:15:38.530 --> 00:15:40.830 So that means that manufacturing needs to start ramping

385 00:15:40.830 --> 00:15:44.841 up right before we ever have a signal of efficacy

386 00:15:44.841 --> 00:15:47.490 and that's a huge risk for companies to take.

387 00:15:47.490 --> 00:15:51.190 So, I'll talk in a couple of slides about sort of how

388 00:15:51.190 --> 00:15:53.630 the government has come in to try to remove

389 00:15:53.630 --> 00:15:56.180 some of that risk from these companies

390 00:15:56.180 --> 00:15:58.670 and then the next slide I think is just showing sort of

391 00:15:58.670 --> 00:16:00.610 that it's really impressive that we're even talking

392 00:16:00.610 --> 00:16:04.770 about potentially having a COVID vaccine available this year

393 00:16:04.770 --> 00:16:07.630 or early next year, just given the timelines

394 00:16:07.630 --> 00:16:10.537 that are required to bring effective vaccines to market.

395 00:16:10.537 --> 00:16:12.600 And so here's just a few, you know,

396 00:16:12.600 --> 00:16:14.160 polio, measles, chickenpox, mumps,

397 00:16:14.160 --> 00:16:18.040 all multiple years of development for these vaccines,

398 00:16:18.040 --> 00:16:19.460 you could add malaria on this list.

399 00:16:19.460 --> 00:16:20.490 It took about 30 years

400 00:16:20.490 --> 00:16:24.210 to get a partially effective malaria vaccine to market.

401 00:16:24.210 --> 00:16:26.373 So this is typically a very long process, right?

402 00:16:26.373 --> 00:16:29.180 And for COVID, we're looking at hopefully doing this

403 00:16:29.180 --> 00:16:31.376 in just under a year or two.

404 00:16:31.376 --> 00:16:34.520 So how is the US government playing a role in this?

405 00:16:34.520 --> 00:16:37.050 Well, it's through this program that you may have heard of

406 00:16:37.050 --> 00:16:39.020 called Operation Warp Speed,

407 00:16:39.020 --> 00:16:43.840 which is this huge convoluted mess of an amalgamation

408 00:16:43.840 --> 00:16:45.350 of programs across the government

409 00:16:45.350 --> 00:16:49.810 from DOD to many branches of NIH, BARDA, NIAID,

410 00:16:49.810 --> 00:16:51.470 so it's sort of all over the place.

411 00:16:51.470 --> 00:16:54.490 And this is really just the same figure

412 00:16:54.490 --> 00:16:56.946 that I showed you from the New England journal paper.

413 00:16:56.946 --> 00:17:00.940 Just maybe a slightly more confusing

414 00:17:00.940 --> 00:17:02.770 if you ask me, I don't think Edward Tufte,

415 00:17:02.770 --> 00:17:04.630 he would be a big fan of graphic

416 00:17:04.630 --> 00:17:07.000 but the point here I want to mention

417 00:17:07.000 --> 00:17:09.500 is how is the government responding

418 00:17:09.500 --> 00:17:10.990 to COVID vaccine development?

419 00:17:10.990 --> 00:17:12.650 How are they contributing to that process?

420 00:17:12.650 --> 00:17:14.640 Well, there's really two ways that they've offered

421 00:17:14.640 --> 00:17:16.590 to accelerate the process.

422 00:17:16.590 --> 00:17:18.910 The first is through funding

423 00:17:18.910 --> 00:17:21.140 of phase three clinical trials, right?

424 00:17:21.140 --> 00:17:23.750 So a number of companies, six of the major companies,

425 00:17:23.750 --> 00:17:25.960 basically every company that's running a phase three trial

426 00:17:25.960 --> 00:17:29.630 in the US besides Pfizer that you've heard about

427 00:17:29.630 --> 00:17:31.640 is contracting with BARDA.

428 00:17:31.640 --> 00:17:33.570 That's an arm of the NIH,

429 00:17:33.570 --> 00:17:36.060 they're contracting with the government

430 00:17:36.060 --> 00:17:38.580 to have the government fund their phase three trials.

431 00:17:38.580 --> 00:17:40.450 So it's a joint agreement between the government

432 00:17:40.450 --> 00:17:42.140 and these companies where the government,

433 00:17:42.140 --> 00:17:44.930 you the taxpayer, right, are paying for these

434 00:17:44.930 --> 00:17:48.550 phase three trials that will eventually lead to licensure.

435 00:17:48.550 --> 00:17:50.300 So that's the first way that the government

436 00:17:50.300 --> 00:17:52.464 is sort of throwing money at this problem.

437 00:17:52.464 --> 00:17:55.850 It's through design and paying for these phase three trials.

438 00:17:55.850 --> 00:17:57.640 The second way is that they're paying

439 00:17:57.640 --> 00:17:58.730 for manufacturing, right?

440 00:17:58.730 --> 00:18:01.120 They're removing that risk for these companies

441 00:18:01.120 --> 00:18:03.810 by basically committing to buy a certain number of doses

442 00:18:03.810 --> 00:18:05.810 before we ever have any efficacy data.

443 00:18:05.810 --> 00:18:08.025 So we're in the hole basically to all of these companies

444 00:18:08.025 --> 00:18:10.540 for a fixed number of doses right.

445 00:18:10.540 --> 00:18:12.900 But that motivates the companies then to scale up

446 00:18:12.900 --> 00:18:14.813 their manufacturing ahead of the time

447 00:18:14.813 --> 00:18:16.693 that efficacy data are available.

448 00:18:17.527 --> 00:18:19.510 And that type of agreement has been entered

449 00:18:19.510 --> 00:18:20.900 into with Pfizer as well.

450 00:18:20.900 --> 00:18:24.460 So all of these companies that OWS Operation Warp Speed

451 00:18:24.460 --> 00:18:26.550 is running the phase three trials for

452 00:18:26.550 --> 00:18:28.980 also have this manufacturing agreement.

453 00:18:28.980 --> 00:18:31.380 Pfizer has that manufacturing agreement as well.

454 00:18:33.130 --> 00:18:37.730 So what role have I played in any of this big messy thing?

455 00:18:37.730 --> 00:18:40.580 So I work with a great group of scientists

456 00:18:40.580 --> 00:18:42.520 in the COVID-19 Prevention Network.

457 00:18:42.520 --> 00:18:45.450 So this was a clinical trials network established

458 00:18:45.450 --> 00:18:48.056 by National Institute of Allergies and Infectious Disease

459 00:18:48.056 --> 00:18:50.780 and NIAID so that's an arm of NIH,

460 00:18:50.780 --> 00:18:54.060 and it's basically anyone who works in clinical trials

461 00:18:54.060 --> 00:18:55.010 is fairly familiar

462 00:18:55.010 --> 00:18:56.840 with these clinical trials networks, right?

463 00:18:56.840 --> 00:19:00.700 It's an amalgamation of researchers and study sites,

464 00:19:00.700 --> 00:19:04.390 laboratories, people who focus on recruitment and retention

465 00:19:04.390 --> 00:19:06.720 of trial participants, statisticians.

466 00:19:06.720 --> 00:19:08.750 So it's researchers who are really experts

467 00:19:08.750 --> 00:19:10.270 in running clinical trials,

468 00:19:10.270 --> 00:19:12.290 designing clinical trials

469 00:19:12.290 --> 00:19:15.070 and ensuring their robust conduct.

470 00:19:15.070 --> 00:19:18.930 So the CoVPN was formed by basically leveraging

471 00:19:18.930 --> 00:19:20.590 four existing clinical trials networks.

472 00:19:20.590 --> 00:19:21.770 One of which I was a part of,

473 00:19:21.770 --> 00:19:23.820 which is the HIV vaccine trials network.

474 00:19:23.820 --> 00:19:27.040 And so from our group, we've really brought a great group

475 00:19:27.040 --> 00:19:29.700 of statisticians, many of whom are at the Fred Hutch

476 00:19:29.700 --> 00:19:34.480 in Seattle as well as great groups of laboratories at U Dub.

477 00:19:34.480 --> 00:19:35.810 And so what are the roles

478 00:19:35.810 --> 00:19:37.530 that we're playing in these trials?

479 00:19:37.530 --> 00:19:40.117 So in our statistical group,

480 00:19:40.117 --> 00:19:44.330 there's a couple of statisticians who are designated

481 00:19:44.330 --> 00:19:46.320 as like CoVPN representatives

482 00:19:46.320 --> 00:19:47.880 for each of these phase three trials.

483 00:19:47.880 --> 00:19:52.680 So I sit on calls with these trials and advise

484 00:19:52.680 --> 00:19:54.950 on their design and analysis approaches

485 00:19:54.950 --> 00:19:57.610 for their efficacy monitoring, for their safety monitoring.

486 00:19:57.610 --> 00:20:01.114 We help them address DSMB and FDA comments

487 00:20:01.114 --> 00:20:03.660 and sort of that's all happening in conjunction

488 00:20:03.660 --> 00:20:06.250 with both government statisticians, right.

489 00:20:06.250 --> 00:20:09.170 Representatives of BARDA and NIAID

490 00:20:10.010 --> 00:20:11.930 as well as company statisticians.

491 00:20:11.930 --> 00:20:14.470 And so we get on these calls and, you know,

492 00:20:14.470 --> 00:20:16.450 nerd out over clinical trials,

493 00:20:16.450 --> 00:20:20.630 statistical decision-making, and it's a good old time.

494 00:20:20.630 --> 00:20:23.660 Another aspect that we really contribute a lot on,

495 00:20:23.660 --> 00:20:26.170 or that CoVPN has sort of been tasked with taking

496 00:20:26.170 --> 00:20:28.870 the lead on is the development of immune correlates.

497 00:20:28.870 --> 00:20:30.590 And so that's the part of my talk

498 00:20:30.590 --> 00:20:32.100 where I'll get a little bit into statistics

499 00:20:32.100 --> 00:20:34.060 and talking about what immune correlates are,

500 00:20:34.060 --> 00:20:35.720 some of the types of analytic approaches

501 00:20:35.720 --> 00:20:38.253 we use to study those and the idea of immune correlates

502 00:20:38.253 --> 00:20:39.880 just to give you a teaser

503 00:20:39.880 --> 00:20:41.710 so you don't, you know, sign off Zoom early.

504 00:20:41.710 --> 00:20:45.430 So immune correlates are really the idea there is

505 00:20:45.430 --> 00:20:48.360 we're looking for immune responses that are predictive

506 00:20:48.360 --> 00:20:51.550 of the vaccines working, right.

507 00:20:51.550 --> 00:20:54.270 So what we'd really like to be able to do is understand,

508 00:20:54.270 --> 00:20:56.040 okay, if we're able to generate this level

509 00:20:56.040 --> 00:20:57.810 of neutralizing antibody,

510 00:20:57.810 --> 00:21:00.174 then that will lead to this level of protective effect

511 00:21:00.174 --> 00:21:01.810 of the vaccine, right?

512 00:21:01.810 --> 00:21:03.960 So that's the whole goal there is identifying

513 00:21:03.960 --> 00:21:05.441 what are these immune responses that are

514 00:21:05.441 --> 00:21:08.443 responsible for providing protection?

515 00:21:08.443 --> 00:21:11.410 Okay so I'm gonna walk through just a few of the design

516 00:21:11.410 --> 00:21:12.243 and analysis questions.

517 00:21:12.243 --> 00:21:14.160 And so these are things that have come up

518 00:21:14.160 --> 00:21:15.990 as we've worked with these company statisticians,

519 00:21:15.990 --> 00:21:19.630 as we thought about sort of the whole OWS vaccine program,

520 00:21:19.630 --> 00:21:21.860 what are some of the issues that statisticians

521 00:21:21.860 --> 00:21:24.120 are kicking around and people who have worked

522 00:21:24.120 --> 00:21:24.990 on clinical trials, right,

523 00:21:24.990 --> 00:21:27.140 a lot of these issues aren't gonna be new

524 00:21:27.140 --> 00:21:30.330 and one thing that I think is sort of interesting about this

525 00:21:30.330 --> 00:21:33.670 whole pandemic and operating as a public health professional

526 00:21:33.670 --> 00:21:36.974 in this and a clinical trial statistician in particular,

527 00:21:36.974 --> 00:21:39.130 is that a lot of things that we take for granted

528 00:21:39.130 --> 00:21:42.210 as scientists are either very confusing

529 00:21:42.210 --> 00:21:44.947 or sort of counterintuitive for a lot of the lay public.

530 00:21:44.947 --> 00:21:48.014 And so it's been sort of interesting to have that laid bare.

531 00:21:48.014 --> 00:21:49.810 In some of these issues, some of these things

532 00:21:49.810 --> 00:21:52.620 that we think are no-brainers like doing interim analysis

533 00:21:52.620 --> 00:21:55.760 for example are kind of highly controversial

534 00:21:55.760 --> 00:21:57.060 and have ended up being, you know,

535 00:21:57.060 --> 00:21:59.350 sort of areas of huge disputes.

536 00:21:59.350 --> 00:22:01.380 And so I just want to run through some of these issues

537 00:22:01.380 --> 00:22:03.730 that I think are quite fascinating, a lot of which,

538 00:22:03.730 --> 00:22:05.880 you know, really don't have a correct answer

539 00:22:05.880 --> 00:22:07.330 and they're really just sort of food for thought

540 00:22:07.330 --> 00:22:09.153 the types of things that we're thinking about

541 00:22:09.153 --> 00:22:11.460 when we're designing these trials.

542 00:22:11.460 --> 00:22:15.620 So I'll start by just giving a sort of more specific idea

543 00:22:15.620 --> 00:22:17.870 of what these trials look like and how they're conducted

544 00:22:17.870 --> 00:22:19.900 and I've picked AstraZeneca because that's the one

545 00:22:19.900 --> 00:22:22.600 I've worked on for the longest and most closely,

546 00:22:22.600 --> 00:22:25.886 but all of the trials sort of follow this similar design.

547 00:22:25.886 --> 00:22:27.080 And so the first thing I'll note

548 00:22:27.080 --> 00:22:28.660 is that you can read these trial protocols.

549 00:22:28.660 --> 00:22:31.230 So one of the interesting things that's happened

550 00:22:31.230 --> 00:22:33.370 in this COVID-19 development processes

551 00:22:33.370 --> 00:22:36.400 is there was a huge public push led by like Eric Topol

552 00:22:36.400 --> 00:22:39.420 and others to have the protocols of these trials

553 00:22:39.420 --> 00:22:42.920 made public, which when it happened was I guess

554 00:22:42.920 --> 00:22:45.220 when that push started happening, you know,

555 00:22:45.220 --> 00:22:46.660 I emailed all my colleagues and said,

556 00:22:46.660 --> 00:22:49.860 really do we not usually make protocols public?

557 00:22:49.860 --> 00:22:51.490 And that was just sort of interesting disconnect

558 00:22:51.490 --> 00:22:53.860 for me as an academic who's used to sort of everything

559 00:22:53.860 --> 00:22:56.920 being open science and that's a no brainer right.

560 00:22:56.920 --> 00:22:58.050 Working in this setting, right,

561 00:22:58.050 --> 00:23:00.489 where these protocols are really seen as trade secrets

562 00:23:00.489 --> 00:23:01.940 for pharmaceutical companies.

563 00:23:01.940 --> 00:23:04.790 So it's really unusual that actually these protocols

564 00:23:04.790 --> 00:23:06.300 for clinical trials have been made public.

565 00:23:06.300 --> 00:23:09.150 So it's sort of neat, but one of the things that happened

566 00:23:09.150 --> 00:23:12.010 is all of these protocols went public and reporters

567 00:23:12.010 --> 00:23:13.380 got their hands on them and said, wow,

568 00:23:13.380 --> 00:23:15.440 these are really dense documents, right?

569 00:23:15.440 --> 00:23:17.628 If you've ever looked at the clinical trial protocol,

570 00:23:17.628 --> 00:23:21.460 it's like a hundred pages of very specific definitions

571 00:23:21.460 --> 00:23:23.720 and safety monitoring and what symptoms lists

572 00:23:23.720 --> 00:23:25.840 you're gonna use and what surveys

573 00:23:25.840 --> 00:23:26.673 you're gonna give to people.

574 00:23:26.673 --> 00:23:28.350 So they're very sort of detailed documents

575 00:23:28.350 --> 00:23:31.810 that are kind of hard for the public to parse.

576 00:23:31.810 --> 00:23:34.430 So it's been sort of a be careful what you wish for thing

577 00:23:34.430 --> 00:23:37.530 in terms of releasing these protocols, but that's an aside.

578 00:23:37.530 --> 00:23:40.030 So let's talk about actually what these trials look like.

579 00:23:40.030 --> 00:23:41.560 So here's a schematic, and again,

580 00:23:41.560 --> 00:23:43.750 this is AstraZeneca in particular,

581 00:23:43.750 --> 00:23:47.230 but this is basically the design of most of these trials

582 00:23:47.230 --> 00:23:48.390 will look something like this.

583 00:23:48.390 --> 00:23:50.020 So who is the population?

584 00:23:50.020 --> 00:23:52.840 Most of these trials are gonna be primarily in adults.

585 00:23:52.840 --> 00:23:54.520 I think Pfizer has now started

586 00:23:54.520 --> 00:23:56.930 to talk about including children.

587 00:23:56.930 --> 00:23:58.684 I'm not exactly sure where that's happening,

588 00:23:58.684 --> 00:24:00.952 but adults for the most part,

589 00:24:00.952 --> 00:24:04.810 these are mostly healthy individuals

590 00:24:04.810 --> 00:24:07.310 that don't have, you know, chronic diseases

591 00:24:07.310 --> 00:24:09.970 that are at risk or high risk of death.

592 00:24:09.970 --> 00:24:11.890 And we're really looking at targeting individuals

593 00:24:11.890 --> 00:24:15.300 who are at an increased risk for SARS-CoV-2 acquisition

594 00:24:15.300 --> 00:24:17.150 and severe COVID disease

595 00:24:17.150 --> 00:24:19.100 and so the idea there is number one

596 00:24:19.100 --> 00:24:20.420 these are the people that are bearing

597 00:24:20.420 --> 00:24:22.520 the brunt of the pandemic, right?

598 00:24:22.520 --> 00:24:25.500 So we want to be able to get a product to those people

599 00:24:25.500 --> 00:24:26.520 as fast as possible.

600 00:24:26.520 --> 00:24:28.603 But number two also, right, that means that we'll accrure

601 00:24:28.603 --> 00:24:32.070 from a sort of cold hearted and statistician point of view

602 00:24:32.070 --> 00:24:34.390 that means we'll accrue end points faster.

603 00:24:34.390 --> 00:24:37.030 We'll observe more cases of COVID-19 disease

604 00:24:37.030 --> 00:24:39.550 and potentially get an efficacy signal a little bit faster.

605 00:24:39.550 --> 00:24:42.500 So there's a lot of interest in sort of recruiting

606 00:24:42.500 --> 00:24:45.050 and retaining individuals at high risk for COVID-19.

607 00:24:45.050 --> 00:24:47.510 So you can go onto the COVID-19 prevention trials network

608 00:24:47.510 --> 00:24:49.050 and fill out a survey, right.

609 00:24:49.050 --> 00:24:50.500 Then we'll basically under the hood

610 00:24:50.500 --> 00:24:52.530 assess your risk for COVID-19

611 00:24:52.530 --> 00:24:53.740 and if you're found to be at high risk,

612 00:24:53.740 --> 00:24:55.400 we'll aggressively email you and try to get you

613 00:24:55.400 --> 00:24:56.580 enrolled in one of these trials.

614 00:24:56.580 --> 00:24:57.413 If you're at low risk,

615 00:24:57.413 --> 00:24:59.040 we'll say, thanks for taking the survey,

616 00:24:59.040 --> 00:25:00.610 we'll be in touch and likely

617 00:25:00.610 --> 00:25:03.474 you won't hear from us anytime soon.

618 00:25:03.474 --> 00:25:05.390 Okay so that's the trial population.

619 00:25:05.390 --> 00:25:07.160 So how does the actual trial conduct look?

620 00:25:07.160 --> 00:25:09.830 So there's kind of a mixture here.

621 00:25:09.830 --> 00:25:12.630 AstraZeneca is using a two to one randomization scheme.

622 00:25:12.630 --> 00:25:15.928 So you have two chances of getting the active vaccine

623 00:25:15.928 --> 00:25:18.190 versus one chance of getting a placebo.

624 00:25:18.190 --> 00:25:21.500 And in this case, it's a true placebo, just a saline dose

625 00:25:21.500 --> 00:25:25.570 and then most of the vaccines, most all with Janssen

626 00:25:25.570 --> 00:25:27.990 being the accepted are two dose vaccines.

627 00:25:27.990 --> 00:25:29.770 So you receive the first dose at day one

628 00:25:29.770 --> 00:25:32.270 and the second dose about a month later.

629 00:25:32.270 --> 00:25:34.110 And in the interim, we take a couple of measurements.

630 00:25:34.110 --> 00:25:36.710 We have a phone call to assess reactogenicity right.

631 00:25:37.964 --> 00:25:41.080 Does your arm hurt, or have you experienced any adverse side

632 00:25:41.080 --> 00:25:44.010 effects of the first dose of vaccine?

633 00:25:44.010 --> 00:25:46.240 And then there's also an immune response measurement

634 00:25:46.240 --> 00:25:47.550 that happens after a couple of days.

635 00:25:47.550 --> 00:25:49.020 So we get an early signal

636 00:25:49.020 --> 00:25:51.247 of how immunogenetic these vaccines are.

637 00:25:51.247 --> 00:25:53.310 And so then individuals come in for their second dose

638 00:25:53.310 --> 00:25:55.100 of vaccine and it's a similar story, right?

639 00:25:55.100 --> 00:25:56.690 Did you have any reactions?

640 00:25:56.690 --> 00:25:59.440 We measure your immune response and after that,

641 00:25:59.440 --> 00:26:02.000 that's sort of when the clock starts for active follow-ups.

642 00:26:02.000 --> 00:26:06.193 So this day 57, that's two weeks roughly after,

643 00:26:07.450 --> 00:26:08.440 am I doing that math right?

644 00:26:08.440 --> 00:26:11.810 Well, it looks like roughly two weeks after the second dose

645 00:26:11.810 --> 00:26:13.810 of the vaccine is typically when this clock

646 00:26:13.810 --> 00:26:16.690 is gonna start and we're gonna start counting COVID events.

647 00:26:16.690 --> 00:26:21.050 And then it's sort of just the standard sort of game we play

648 00:26:21.050 --> 00:26:21.883 in clinical trials.

649 00:26:21.883 --> 00:26:23.310 We wait for events to accrue.

650 00:26:23.310 --> 00:26:24.930 We have certain monitoring plan

651 00:26:24.930 --> 00:26:26.840 for when we're gonna check for efficacy

652 00:26:26.840 --> 00:26:28.020 and we'll talk about some of that.

653 00:26:28.020 --> 00:26:30.350 So, I just want to note that there's sort of two ways

654 00:26:30.350 --> 00:26:31.590 that we're ascertaining events

655 00:26:31.590 --> 00:26:33.090 that are happening here, right?

656 00:26:33.090 --> 00:26:34.840 The first is passive monitoring.

657 00:26:34.840 --> 00:26:37.030 What that means is we basically wait for individuals

658 00:26:37.030 --> 00:26:39.230 to present with symptoms of COVID 19, right?

659 00:26:39.230 --> 00:26:41.400 So you get a cough, you lose taste, right?

660 00:26:41.400 --> 00:26:44.810 You call the study site, right?

661 00:26:44.810 --> 00:26:45.810 So I am having these symptoms.

662 00:26:45.810 --> 00:26:46.910 They say, come on in.

663 00:26:46.910 --> 00:26:49.642 You get a PCR test to see whether you're infected.

664 00:26:49.642 --> 00:26:51.310 And in that case, you would count

665 00:26:51.310 --> 00:26:52.900 as a COVID-19 endpoint, right?

666 00:26:52.900 --> 00:26:55.700 If you check off some check boxes for symptoms

667 00:26:55.700 --> 00:26:58.570 with COVID-19 disease, you have a PCR positive test.

668 00:26:58.570 --> 00:27:01.100 You'd go down as a COVID 19 endpoint.

669 00:27:01.100 --> 00:27:04.000 There's also these sort of active follow-up visits.

670 00:27:04.000 --> 00:27:07.559 So these like day 90, day, 180 and day 360,

671 00:27:07.559 --> 00:27:10.580 and at those visits we'll do a serology check.

672 00:27:10.580 --> 00:27:12.560 And what that means is we basically take a blood draw

673 00:27:12.560 --> 00:27:15.670 and we measure whether you have antibodies

674 00:27:15.670 --> 00:27:18.640 against SARS-CoV-2, right, antibodies that are distinct

675 00:27:18.640 --> 00:27:19.970 from the antibodies that are generated

676 00:27:19.970 --> 00:27:21.110 in response to the vaccine.

677 00:27:21.110 --> 00:27:24.130 So we're basically able to tell whether you were infected

678 00:27:24.130 --> 00:27:26.270 in this sort of interim period,

679 00:27:26.270 --> 00:27:28.740 when you show up for these visits.

680 00:27:28.740 --> 00:27:30.290 So that's active follow up

681 00:27:30.290 --> 00:27:31.270 and so there you're gonna be able

682 00:27:31.270 --> 00:27:33.390 to pick up sort of asymptomatic cases, right?

683 00:27:33.390 --> 00:27:35.940 'Cause if you never have symptoms, you'll never come in

684 00:27:35.940 --> 00:27:38.040 and be captured by passive followup.

685 00:27:38.040 --> 00:27:40.070 So we have to wait for these set clinic visits

686 00:27:40.070 --> 00:27:41.720 to do the serology testing,

687 00:27:41.720 --> 00:27:43.820 to ascertain it asymptomatic cases.

688 00:27:43.820 --> 00:27:46.010 And so this is gonna actually play a role

689 00:27:46.010 --> 00:27:47.600 in a little bit, when I started talking about, you know,

690 00:27:47.600 --> 00:27:50.110 what are the end points that we're thinking about measuring?

691 00:27:50.110 --> 00:27:51.530 Like, what do we want to know how well

692 00:27:51.530 --> 00:27:53.270 the vaccine works at preventing?

693 00:27:53.270 --> 00:27:54.830 Is it asymptomatic infection?

694 00:27:54.830 --> 00:27:55.990 Is it disease?

695 00:27:55.990 --> 00:27:57.470 Is it severe disease and so forth?

696 00:27:57.470 --> 00:27:58.870 So we'll talk through some of those issues,

697 00:27:58.870 --> 00:28:01.720 but just want to note already that the design has started

698 00:28:01.720 --> 00:28:04.350 to inform some of the challenges that we might see

699 00:28:04.350 --> 00:28:06.510 when we want to talk about how well the vaccine works

700 00:28:06.510 --> 00:28:09.113 against certain forms of infection and disease.

701 00:28:10.220 --> 00:28:12.970 And so I think if you read the newspaper and you'll see

702 00:28:12.970 --> 00:28:15.290 the term vaccine efficacy tossed around a lot.

703 00:28:15.290 --> 00:28:16.830 So the first thing I want to talk about is right,

704 00:28:16.830 --> 00:28:19.080 what is the primary hypothesis

705 00:28:19.080 --> 00:28:20.530 that these trials are trying to test?

706 00:28:20.530 --> 00:28:22.510 And what is the parameter?

707 00:28:22.510 --> 00:28:24.820 What is the estimate, right, that they're going after

708 00:28:24.820 --> 00:28:26.550 in these trials and for whatever reason

709 00:28:26.550 --> 00:28:28.800 nobody consulted me when they decided that VE

710 00:28:28.800 --> 00:28:30.933 would be measured in this way.

711 00:28:31.782 --> 00:28:35.220 But for whatever reason, we studied this that we quantify

712 00:28:35.220 --> 00:28:37.180 the efficacy of a vaccine in a sort of weird way.

713 00:28:37.180 --> 00:28:40.300 So a vaccine efficacy, we describe as the percent reduction

714 00:28:40.300 --> 00:28:43.040 in relative risk comparing vaccine to placebo.

715 00:28:43.040 --> 00:28:46.060 So it's this one minus a risk ratio.

716 00:28:46.060 --> 00:28:49.140 There's a one minus a risk ratio where you take the risk

717 00:28:49.140 --> 00:28:51.300 in the vaccine and the numerator and the risk

718 00:28:51.300 --> 00:28:53.410 in the placebo and the denominator.

719 00:28:53.410 --> 00:28:54.243 So, I mean,

720 00:28:54.243 --> 00:28:55.940 we can just play a quick little intuitive game, right?

721 00:28:55.940 --> 00:28:57.280 How do we get a VE close

722 00:28:57.280 --> 00:28:59.650 to one that would be a perfect vaccine?

723 00:28:59.650 --> 00:29:00.720 Well, we would make the risk

724 00:29:00.720 --> 00:29:02.840 in the vaccine close to zero, right?

725 00:29:02.840 --> 00:29:03.780 So that sorta makes sense.

726 00:29:03.780 --> 00:29:05.640 If you have a perfectly effective vaccine,

727 00:29:05.640 --> 00:29:08.350 there'll be no risk of infection and or disease

728 00:29:08.350 --> 00:29:09.310 amongst the vaccinated.

729 00:29:09.310 --> 00:29:11.210 So you would get VE close to one.

730 00:29:11.210 --> 00:29:13.498 But on the other hand, how do we make VE zero?

731 00:29:13.498 --> 00:29:15.660 Well, we would take the risk in the vaccine

732 00:29:15.660 --> 00:29:18.050 and set it equal to the risk in the placebo, right.

733 00:29:18.050 --> 00:29:20.680 In which case basically saying the vaccine's not doing

734 00:29:20.680 --> 00:29:22.620 anything and then on the other hand,

735 00:29:22.620 --> 00:29:23.860 a VE is negative, right?

736 00:29:23.860 --> 00:29:26.040 That's indicating that there's actually higher risk

737 00:29:26.040 --> 00:29:27.890 in the vaccine.

738 00:29:27.890 --> 00:29:30.840 So just to give you sort of a few reference points, right?

739 00:29:30.840 --> 00:29:34.200 So that VE of one is perfect, VE of zero is nothing

740 00:29:34.200 --> 00:29:37.590 and what we're really hoping for with these COVID trials

741 00:29:37.590 --> 00:29:39.730 is a VE of at least 50%.

742 00:29:39.730 --> 00:29:42.460 And that's sort of the cutoff that FDA guidance

743 00:29:42.460 --> 00:29:45.590 has stipulated is that you need to show a point estimate

744 00:29:45.590 --> 00:29:48.050 of VE for your primary end point.

745 00:29:48.050 --> 00:29:50.040 And again, we'll talk about what these primary end points

746 00:29:50.040 --> 00:29:52.620 are but we need a VE against a primary end point

747 00:29:52.620 --> 00:29:54.147 of at least 50%

748 00:29:54.147 --> 00:29:59.147 and we need to definitively rule out the possibility

749 00:29:59.480 --> 00:30:02.250 that the vaccine efficacy is less than 30%.

750 00:30:02.250 --> 00:30:04.500 So basically we have to reject the null hypothesis

751 00:30:04.500 --> 00:30:07.710 that VE is less than 30% along with having a point

752 00:30:07.710 --> 00:30:10.760 estimate of VE being greater than 50%, right.

753 00:30:10.760 --> 00:30:13.273 And we need to do that while controlling type one error

754 00:30:13.273 --> 00:30:14.863 at two and a half percent.

755 00:30:15.820 --> 00:30:18.620 Okay and so here, just one final note,

756 00:30:18.620 --> 00:30:19.990 since this is a statistics talk,

757 00:30:19.990 --> 00:30:21.210 I'll talk a little bit more

758 00:30:21.210 --> 00:30:23.110 about what I mean by risk, right?

759 00:30:23.110 --> 00:30:25.730 So risk here can be quantified in a number of ways

760 00:30:25.730 --> 00:30:26.563 and it often is.

761 00:30:26.563 --> 00:30:28.920 So we can quantify this using hazards, for example,

762 00:30:28.920 --> 00:30:31.210 like you can imagine fitting a Cox model, right.

763 00:30:31.210 --> 00:30:32.580 A proportional hazards model, right.

764 00:30:32.580 --> 00:30:34.720 That only adjusts for vaccine, right.

765 00:30:34.720 --> 00:30:36.880 And presenting like one minus a hazard ratio

766 00:30:36.880 --> 00:30:39.250 from a Cox model, that's something that's commonly done.

767 00:30:39.250 --> 00:30:41.430 You can also think about cumulative incidents, right?

768 00:30:41.430 --> 00:30:42.720 So like mapping,

769 00:30:42.720 --> 00:30:45.980 maybe one minus a survival probability as a way

770 00:30:45.980 --> 00:30:49.140 of quantifying risk, incidents rate ratios.

771 00:30:49.140 --> 00:30:51.990 So they're all sort of used for different vaccines.

772 00:30:51.990 --> 00:30:54.530 And usually we like to sort of argue

773 00:30:54.530 --> 00:30:55.890 about which one of these is better

774 00:30:55.890 --> 00:30:58.390 and I've thought a lot about that in my career.

775 00:30:58.390 --> 00:31:00.140 And in this setting, it turns out because COVID

776 00:31:00.140 --> 00:31:02.930 is such a rare event that all of these ways of quantifying

777 00:31:02.930 --> 00:31:05.220 rates are basically the same and you end up

778 00:31:05.220 --> 00:31:07.960 with almost identical operating characteristics of a trial.

779 00:31:07.960 --> 00:31:09.840 So it's really not worth sort of losing sleep over

780 00:31:09.840 --> 00:31:12.107 whether we're talking about VE in terms of hazard

781 00:31:12.107 --> 00:31:14.563 or incidents or incidents rate and so forth.

782 00:31:16.470 --> 00:31:18.830 So how are folks going about estimating this VE?

783 00:31:18.830 --> 00:31:21.950 Here's just a quick table of the four most advanced

784 00:31:21.950 --> 00:31:22.783 phase three trials,

785 00:31:22.783 --> 00:31:25.330 the ones that have released their protocols at least.

786 00:31:25.330 --> 00:31:28.200 So we see for Moderna, AstraZeneca, and Janssen,

787 00:31:28.200 --> 00:31:30.800 they're using pretty kind of the standard approaches.

788 00:31:30.800 --> 00:31:33.140 Moderna a Cox model as I describe,

789 00:31:33.140 --> 00:31:35.170 AstraZeneca a Poisson regression model,

790 00:31:35.170 --> 00:31:37.810 it's like, okay, that's basically a Cox model,

791 00:31:37.810 --> 00:31:40.660 and then Janssen is using a sort of exact binomial test

792 00:31:40.660 --> 00:31:43.681 with this sequential probability ratio rest.

793 00:31:43.681 --> 00:31:46.320 Pfizer is a little bit of the oddball.

794 00:31:46.320 --> 00:31:49.030 So they have stipulated a bayesian approach

795 00:31:49.030 --> 00:31:52.770 wherein they're basically specifying a prior

796 00:31:52.770 --> 00:31:55.290 for vaccine efficacy and are using sort of

797 00:31:55.290 --> 00:31:57.880 a beta-binomial bayesian approach to evaluate

798 00:31:57.880 --> 00:31:59.840 the posterior probability of the vaccine efficacy

799 00:31:59.840 --> 00:32:03.970 is greater than 30% and so at the end of the day,

800 00:32:03.970 --> 00:32:05.530 there's four different statistical methods here.

801 00:32:05.530 --> 00:32:08.850 Again, if you do a simulation study with parameters

802 00:32:08.850 --> 00:32:10.760 that are approximately similar to what we expect to see

803 00:32:10.760 --> 00:32:12.010 in these COVID trials,

804 00:32:12.010 --> 00:32:13.930 you're really not gonna see much difference in terms

805 00:32:13.930 --> 00:32:15.740 of operating characteristics across these.

806 00:32:15.740 --> 00:32:17.850 So it's interesting to notice that assertions

807 00:32:17.850 --> 00:32:19.410 that there's these different approaches,

808 00:32:19.410 --> 00:32:20.243 but at the end of the day,

809 00:32:20.243 --> 00:32:22.440 we're basically talking about how many vaccinated people

810 00:32:22.440 --> 00:32:25.080 get infected, how many placebo people got infected,

811 00:32:25.080 --> 00:32:27.410 and almost all of these methods are gonna yield

812 00:32:27.410 --> 00:32:29.010 very similar inference.

813 00:32:29.010 --> 00:32:30.500 When it comes down to brass tacks,

814 00:32:30.500 --> 00:32:33.190 how many numbers fall into those categories?

815 00:32:33.190 --> 00:32:35.710 So that's a little bit about sort of

816 00:32:35.710 --> 00:32:38.080 how we quantify VE in these settings

817 00:32:38.080 --> 00:32:39.920 but one of the big things I haven't described yet

818 00:32:39.920 --> 00:32:41.880 is VE against what, right?

819 00:32:41.880 --> 00:32:43.497 What is the end point that we're measuring here?

820 00:32:43.497 --> 00:32:47.280 And so here's a figure from a paper we just had come out

821 00:32:47.280 --> 00:32:50.040 in Annals of Internal Medicine, the link's here.

822 00:32:50.040 --> 00:32:52.080 So this is where we were spending a lot of time,

823 00:32:52.080 --> 00:32:54.230 you know, earlier this summer, thinking about,

824 00:32:54.230 --> 00:32:55.920 you know, what's the right end point,

825 00:32:55.920 --> 00:32:57.900 what's the right end point for a primary analysis

826 00:32:57.900 --> 00:32:59.257 of the clinical trial.

827 00:32:59.257 --> 00:33:02.530 And it's complicated for something like SARS-CoV-2, right?

828 00:33:02.530 --> 00:33:03.733 Because we know we can start up here

829 00:33:03.733 --> 00:33:07.030 with the SARS-CoV-2 infection, right?

830 00:33:07.030 --> 00:33:08.590 That's sort of the base, you can become infected

831 00:33:08.590 --> 00:33:10.930 and then a number of things can happen, right?

832 00:33:10.930 --> 00:33:14.270 You can go on to be infected but develop no symptoms.

833 00:33:14.270 --> 00:33:16.750 So we would call that an asymptomatic infection,

834 00:33:16.750 --> 00:33:18.410 or you can develop symptoms, right.

835 00:33:18.410 --> 00:33:19.700 In which case we don't call you

836 00:33:19.700 --> 00:33:21.300 a SAR-CoV-2 infection anymore,

837 00:33:21.300 --> 00:33:24.880 we call you a COVID-19 disease endpoint.

838 00:33:24.880 --> 00:33:28.480 You have a clinical manifestation of your infection.

839 00:33:28.480 --> 00:33:29.640 But even beyond that, right,

840 00:33:29.640 --> 00:33:31.220 amongst people who exhibit symptoms

841 00:33:31.220 --> 00:33:34.380 some of them, maybe many of them are quite mild, right.

842 00:33:34.380 --> 00:33:37.580 So we have this kind of category of non-severe COVID,

843 00:33:37.580 --> 00:33:40.730 whereas others we know that are extremely adversely

844 00:33:40.730 --> 00:33:45.330 impacted by infection and end up with severe COVID disease.

845 00:33:45.330 --> 00:33:48.740 So you have all of these choices of sort of

846 00:33:48.740 --> 00:33:50.820 which end points you might want to talk about

847 00:33:50.820 --> 00:33:53.050 and so I'll kind of walk through some what I see

848 00:33:53.050 --> 00:33:56.360 as the positives and negatives of this and then I'll also

849 00:33:56.360 --> 00:33:57.820 talk about this burden of disease

850 00:33:57.820 --> 00:34:00.907 very briefly end point that we've put together

851 00:34:00.907 --> 00:34:02.780 and so that's kind of a composite end point

852 00:34:02.780 --> 00:34:05.000 that we've suggested that could kind of bring all

853 00:34:05.000 --> 00:34:07.170 of these different end points together.

854 00:34:07.170 --> 00:34:09.360 Okay so starting with SARS-CoV-2 infection, right?

855 00:34:09.360 --> 00:34:12.390 Why might we like any sort of any infection, right.

856 00:34:12.390 --> 00:34:14.020 Asymptomatic, symptomatic don't care,

857 00:34:14.020 --> 00:34:16.730 let's count any infection as an event

858 00:34:16.730 --> 00:34:19.660 and measure VE against preventing infection.

859 00:34:19.660 --> 00:34:21.620 Okay and so that's definitely relevant, right.

860 00:34:21.620 --> 00:34:23.940 It's relevant the context of a pandemic.

861 00:34:23.940 --> 00:34:24.930 We're preventing infections,

862 00:34:24.930 --> 00:34:26.520 we're preventing spread of the disease,

863 00:34:26.520 --> 00:34:29.967 we're bringing our knot down, we're impacting the pandemic.

864 00:34:29.967 --> 00:34:33.080 And moreover, we're going to see many more infections

865 00:34:33.080 --> 00:34:35.580 than we will cases of symptomatic disease.

866 00:34:35.580 --> 00:34:37.220 We know that many people who were infected

867 00:34:37.220 --> 00:34:39.110 never go on to develop symptoms

868 00:34:39.110 --> 00:34:42.310 so thinking about having an answer faster, right.

869 00:34:42.310 --> 00:34:44.469 SARS-CoV-2 infection is a nice endpoint,

870 00:34:44.469 --> 00:34:45.780 but then the question is,

871 00:34:45.780 --> 00:34:47.430 is it a clinically relevant endpoint?

872 00:34:47.430 --> 00:34:52.085 So it's really not describing an impact on patients at all.

873 00:34:52.085 --> 00:34:55.510 So we could kind of question its relevance

874 00:34:55.510 --> 00:34:56.940 from that perspective.

875 00:34:56.940 --> 00:34:58.710 The other thing, right, is that we remember going back

876 00:34:58.710 --> 00:35:01.080 to the study design, we're only able to ascertain

877 00:35:01.080 --> 00:35:05.270 asymptomatic infections sort of very coarsely in time

878 00:35:05.270 --> 00:35:08.930 and moreover you have this phenomenon that happens

879 00:35:08.930 --> 00:35:12.160 is that when you're testing many, many individuals, right.

880 00:35:12.160 --> 00:35:13.587 It's sort of the classic biostat

881 00:35:13.587 --> 00:35:15.530 one-on-one problem that we give people, right.

882 00:35:15.530 --> 00:35:18.870 You're testing many individuals, but the prevalence is low.

883 00:35:18.870 --> 00:35:22.470 So even if you have high sensitivity and high specificity,

884 00:35:22.470 --> 00:35:24.620 you could end up with low positive predictive value.

885 00:35:24.620 --> 00:35:28.480 And the effect of that when you come to the time to analyze

886 00:35:28.480 --> 00:35:31.320 the data is that you'll be biasing VE towards the knoll.

887 00:35:31.320 --> 00:35:35.065 So it's actually, while it seems like maybe a nice end point

888 00:35:35.065 --> 00:35:36.920 from the perspective of observing many infections,

889 00:35:36.920 --> 00:35:40.343 it's a very challenging endpoint to analyze quantitatively.

890 00:35:41.190 --> 00:35:43.260 So moving down we could talk about COVID.

891 00:35:43.260 --> 00:35:45.690 So again, COVID is just infection,

892 00:35:45.690 --> 00:35:48.880 PCR confirmed infection with clinical symptoms.

893 00:35:48.880 --> 00:35:50.770 So that's of course more clinically relevant, right.

894 00:35:50.770 --> 00:35:52.560 Because we're starting to talk about

895 00:35:52.560 --> 00:35:56.830 an impact, excuse me, the endpoint that impacts patients.

896 00:35:56.830 --> 00:36:00.090 All right so that's more clinically relevant and moreover

897 00:36:00.090 --> 00:36:03.130 we'll expect to have a reasonable number of cases, right.

898 00:36:03.130 --> 00:36:06.410 By including more mild cases, for example,

899 00:36:06.410 --> 00:36:08.590 in this endpoint definition.

900 00:36:08.590 --> 00:36:10.360 But then on the other side of that coin

901 00:36:10.360 --> 00:36:12.280 is it really that clinically relevant

902 00:36:12.280 --> 00:36:14.540 if we're just talking about mild symptoms?

903 00:36:14.540 --> 00:36:16.010 We're talking about a disease where you get it

904 00:36:16.010 --> 00:36:17.930 and you end up with a little cough for a couple of weeks

905 00:36:17.930 --> 00:36:19.350 and that's it.

906 00:36:19.350 --> 00:36:22.020 So then maybe you suggest using severe COVID right.

907 00:36:22.020 --> 00:36:23.940 That's the most clinically relevant one.

908 00:36:23.940 --> 00:36:26.420 We want to be protecting the most vulnerable individuals

909 00:36:26.420 --> 00:36:28.480 so we should be quantifying how well our vaccines

910 00:36:28.480 --> 00:36:33.130 work towards preventing those most severe end points.

911 00:36:33.130 --> 00:36:34.990 And so most clinically relevant,

912 00:36:34.990 --> 00:36:36.930 and also there's sort of a long history

913 00:36:36.930 --> 00:36:40.000 of vaccine development where really we see the best VE

914 00:36:40.000 --> 00:36:42.890 against severe cases of disease.

915 00:36:42.890 --> 00:36:44.740 So that's really where we expect the vaccines

916 00:36:44.740 --> 00:36:47.150 to have the most impact is maybe we are not preventing

917 00:36:47.150 --> 00:36:50.580 you from being infected but we're lessening the symptoms

918 00:36:50.580 --> 00:36:52.410 once you become infected.

919 00:36:52.410 --> 00:36:54.880 So we're not totally blocking transmission

920 00:36:54.880 --> 00:36:56.730 but we're making a clinical impact on disease

921 00:36:56.730 --> 00:36:57.750 and that's sort of been seen

922 00:36:57.750 --> 00:37:00.280 for a number of vaccines in the past.

923 00:37:00.280 --> 00:37:01.670 The downside of this end point of course

924 00:37:01.670 --> 00:37:04.150 is that there's very few cases expected to be observed.

925 00:37:04.150 --> 00:37:05.350 So amongst all infections,

926 00:37:05.350 --> 00:37:07.420 only a fraction have any symptoms.

927 00:37:07.420 --> 00:37:08.640 Amongst those with any symptoms,

928 00:37:08.640 --> 00:37:10.330 only a fraction develops severe symptoms.

929 00:37:10.330 --> 00:37:13.040 So we're really whittling away the number of end points.

930 00:37:13.040 --> 00:37:14.500 So we need to do larger trials

931 00:37:14.500 --> 00:37:17.703 or have longer follow-up to evaluate this endpoint.

932 00:37:18.890 --> 00:37:21.390 And so in that paper, I'm sort of pressed for time

933 00:37:21.390 --> 00:37:23.700 so I won't spend too much time talking about this,

934 00:37:23.700 --> 00:37:26.280 we also proposed this burden of disease measure

935 00:37:26.280 --> 00:37:29.350 where you're sort of scoring these these outcomes, right?

936 00:37:29.350 --> 00:37:31.030 So maybe you would get a score of zero

937 00:37:31.030 --> 00:37:32.710 if you're an asymptomatic infection

938 00:37:32.710 --> 00:37:35.780 'cause it's really no burden on you as a patient, right?

939 00:37:35.780 --> 00:37:37.370 You don't have any symptoms.

940 00:37:37.370 --> 00:37:39.490 And then we're sort of assigning arbitrarily

941 00:37:39.490 --> 00:37:42.020 a score of one for non severe COVID so that's like

942 00:37:42.020 --> 00:37:44.610 mild cases of COVID and a score of two

943 00:37:44.610 --> 00:37:47.990 for severe cases of COVID and this end point actually

944 00:37:47.990 --> 00:37:50.680 has some nice operating characteristics we think,

945 00:37:50.680 --> 00:37:53.160 but of course it's subject to controversy, anytime you start

946 00:37:53.160 --> 00:37:57.370 talking about an ordinal scoring system, right,

947 00:37:57.370 --> 00:37:59.450 you start to raise questions about how you're assigning

948 00:37:59.450 --> 00:38:01.200 the burden of disease score, right?

949 00:38:01.200 --> 00:38:03.430 Why should severe cases be a two

950 00:38:03.430 --> 00:38:05.620 versus a three versus a five and so forth?

951 00:38:05.620 --> 00:38:07.460 So you can kind of get bogged down

952 00:38:07.460 --> 00:38:09.193 in some of the specifics of that.

953 00:38:10.220 --> 00:38:11.820 So what has FDA said about this?

954 00:38:11.820 --> 00:38:14.600 So FDA guidance documents states that either

955 00:38:14.600 --> 00:38:17.830 the COVID end point or SARS-CoV-2 infection

956 00:38:17.830 --> 00:38:19.310 is an acceptable primary endpoint

957 00:38:19.310 --> 00:38:22.180 and then somewhat ironically OWS has been telling companies

958 00:38:22.180 --> 00:38:23.950 that infection alone is not acceptable

959 00:38:23.950 --> 00:38:24.870 as a primary end point.

960 00:38:24.870 --> 00:38:27.590 So we had one company that was interested in including

961 00:38:27.590 --> 00:38:31.150 that as co-primary and for whatever reason we told them

962 00:38:31.150 --> 00:38:36.061 please don't do that, and then beyond that so COVID

963 00:38:36.061 --> 00:38:38.570 has sort of won out as the end point of choice.

964 00:38:38.570 --> 00:38:41.620 But beyond that FDA guidance states that companies should

965 00:38:41.620 --> 00:38:44.060 consider powering efficacy trials

966 00:38:44.060 --> 00:38:48.230 for the severe COVID endpoint as a co-primary or at least

967 00:38:48.230 --> 00:38:50.510 as a key secondary endpoint in the trial.

968 00:38:50.510 --> 00:38:53.690 And so so far only Janssen has taken them up on that offer

969 00:38:53.690 --> 00:38:55.800 of making severe COVID primary.

970 00:38:55.800 --> 00:38:58.010 And that's why, if you look at the number of individuals

971 00:38:58.010 --> 00:38:59.290 that are planning to enroll in their trial,

972 00:38:59.290 --> 00:39:02.530 it's twice as many as any of the other OWS trials.

973 00:39:02.530 --> 00:39:04.443 So like AstraZeneca is planning for 30,000,

974 00:39:04.443 --> 00:39:07.730 Janssen is planning for 60,000 in their trial.

975 00:39:07.730 --> 00:39:10.480 And that's the power, to see more cases of severe disease

976 00:39:10.480 --> 00:39:14.193 to be sufficiently powered to detect VE against that.

977 00:39:15.100 --> 00:39:17.337 So this is a controversial slide.

978 00:39:17.337 --> 00:39:20.330 Or this is virtual topic I found,

979 00:39:20.330 --> 00:39:22.480 again, something that clinical trials statisticians

980 00:39:22.480 --> 00:39:25.570 sort of take for granted is doing interim analyses, right?

981 00:39:25.570 --> 00:39:27.990 If the treatment is working and we have enough evidence

982 00:39:27.990 --> 00:39:29.430 to claim that a treatment is working,

983 00:39:29.430 --> 00:39:31.420 we'd like to stop that trial early

984 00:39:31.420 --> 00:39:32.990 to get that treatment to patients, right.

985 00:39:32.990 --> 00:39:34.830 One would think that that's true here

986 00:39:34.830 --> 00:39:36.340 and so many of these trials

987 00:39:36.340 --> 00:39:40.180 were designed with aggressive sort of interim looks, right?

988 00:39:40.180 --> 00:39:41.460 Because we're in the middle of the pandemic

989 00:39:41.460 --> 00:39:44.070 and we'd like to get a vaccine to individuals

990 00:39:44.070 --> 00:39:44.903 as quickly as possible.

991 00:39:44.903 --> 00:39:47.108 So I have a table, we won't go through it all here,

992 00:39:47.108 --> 00:39:50.370 just sort of the planned interim analysis

993 00:39:50.370 --> 00:39:52.010 for these different trials.

994 00:39:52.010 --> 00:39:55.710 I would say Pfizer seems to be the most aggressive so far.

995 00:39:55.710 --> 00:39:59.680 They have five interim looks or four interim looks

996 00:39:59.680 --> 00:40:01.770 and a final look at their data, right?

997 00:40:01.770 --> 00:40:03.340 So that's fairly aggressive.

998 00:40:03.340 --> 00:40:06.540 OWS again, the trials that we're running,

999 00:40:06.540 --> 00:40:08.850 we're really encouraging companies to be a bit

1000 00:40:08.850 --> 00:40:10.830 more conservative in the approach to this

1001 00:40:10.830 --> 00:40:12.740 and only maybe two or three

1002 00:40:12.740 --> 00:40:14.290 and so you see what's been adopted

1003 00:40:14.290 --> 00:40:17.110 by Moderna and AstraZeneca

1004 00:40:17.110 --> 00:40:19.470 and so this was really a big point of contention

1005 00:40:19.470 --> 00:40:22.270 I think when these protocols were made public is this idea

1006 00:40:22.270 --> 00:40:25.390 that like, can you really know that a vaccine works

1007 00:40:25.390 --> 00:40:27.410 based on 32 data points, right?

1008 00:40:27.410 --> 00:40:30.370 We're talking about a vaccine that's going to be given

1009 00:40:30.370 --> 00:40:31.890 to billions of people around

1010 00:40:31.890 --> 00:40:33.280 the world based on these results

1011 00:40:33.280 --> 00:40:34.720 and you're gonna do that based

1012 00:40:34.720 --> 00:40:37.280 on the results in 32 individuals?

1013 00:40:37.280 --> 00:40:39.720 And like, so I can stare at the math and say that like, yes,

1014 00:40:39.720 --> 00:40:42.420 that appropriately controls type one error and so forth,

1015 00:40:42.420 --> 00:40:44.780 but it still makes me just feel a little bit uncomfortable.

1016 00:40:44.780 --> 00:40:47.500 There's a bit of dissonance between sort of my life

1017 00:40:47.500 --> 00:40:50.210 as a statistician and just me being a human

1018 00:40:50.210 --> 00:40:52.130 and saying 32 data points is probably not enough

1019 00:40:52.130 --> 00:40:54.200 to decide to vaccinate billions of people.

1020 00:40:54.200 --> 00:40:56.509 And so a lot of people I think sort of shared

1021 00:40:56.509 --> 00:41:00.880 that viewpoint and in response FDA has now been sort of

1022 00:41:00.880 --> 00:41:05.630 backpedaling in a way and asking companies to provide more

1023 00:41:05.630 --> 00:41:10.020 data in order to grant an emergency authorization

1024 00:41:10.020 --> 00:41:10.853 for their vaccine.

1025 00:41:10.853 --> 00:41:14.490 So this EUA mechanism that FDA has of approving vaccines.

1026 00:41:14.490 --> 00:41:16.730 And so in addition to an efficacy signal,

1027 00:41:16.730 --> 00:41:19.840 now companies also are gonna be required, I think,

1028 00:41:19.840 --> 00:41:22.650 and this is sort of still a moving target so this is maybe

1029 00:41:22.650 --> 00:41:25.930 like data news at this point but I think prior to offering

1030 00:41:25.930 --> 00:41:29.260 an EUA, FDA has now said that companies need to have 50%

1031 00:41:29.260 --> 00:41:32.511 of participants complete at least two months of follow-up

1032 00:41:32.511 --> 00:41:36.151 for safety signals and that you need to have at least

1033 00:41:36.151 --> 00:41:38.560 six COVID cases in the oldest age group.

1034 00:41:38.560 --> 00:41:40.820 Of course, that's an age group of particular interest

1035 00:41:40.820 --> 00:41:43.720 in terms of severe cases and at least five cases

1036 00:41:43.720 --> 00:41:45.400 of severe COVID in the placebo group.

1037 00:41:45.400 --> 00:41:47.830 So they want to be able to see some data,

1038 00:41:47.830 --> 00:41:50.090 even if you're not specifying severe COVID

1039 00:41:50.090 --> 00:41:51.100 as a primary end point,

1040 00:41:51.100 --> 00:41:52.800 they want to be able to see some data,

1041 00:41:52.800 --> 00:41:54.500 some signal of efficacy against that

1042 00:41:54.500 --> 00:41:55.913 in order to grant licensure.

1043 00:41:56.770 --> 00:42:00.539 So I'll sort of, I won't go through this slide.

1044 00:42:00.539 --> 00:42:01.960 It's just to say that like,

1045 00:42:01.960 --> 00:42:03.980 sort of when Pfizer released their protocol,

1046 00:42:03.980 --> 00:42:06.410 everyone was like, ooh a bayesian analysis

1047 00:42:06.410 --> 00:42:08.760 and got very sort of skeptical, right?

1048 00:42:08.760 --> 00:42:10.530 Because the Pfizer CEO has been out there

1049 00:42:10.530 --> 00:42:12.510 sort of chest thumping and saying they're gonna have

1050 00:42:12.510 --> 00:42:15.040 a vaccine before the election and so forth

1051 00:42:15.040 --> 00:42:16.620 and then they came out with this bayesian design

1052 00:42:16.620 --> 00:42:18.990 that was a little atypical and so everybody was asking

1053 00:42:18.990 --> 00:42:21.210 the question, well, are they trying to hide something?

1054 00:42:21.210 --> 00:42:22.670 So I sort of did a quick analysis

1055 00:42:22.670 --> 00:42:25.120 and found that really it doesn't look that different

1056 00:42:25.120 --> 00:42:28.080 than a classic kind of post hoc monitored design.

1057 00:42:28.080 --> 00:42:29.550 And if you want to read more about that,

1058 00:42:29.550 --> 00:42:32.919 I have some slides up on my GitHub about it.

1059 00:42:32.919 --> 00:42:35.870 So let's see, I'm running low on time

1060 00:42:35.870 --> 00:42:38.760 so I'm gonna skip over sort of the question

1061 00:42:38.760 --> 00:42:40.830 of what happens if efficacy is declared early.

1062 00:42:40.830 --> 00:42:43.330 So I have some reasons that we should be excited, right?

1063 00:42:43.330 --> 00:42:45.537 If one of these trials stops earlier, I can get a vaccine.

1064 00:42:45.537 --> 00:42:48.460 There's good data that the vaccine works

1065 00:42:48.460 --> 00:42:49.790 and that's nice.

1066 00:42:49.790 --> 00:42:52.310 I'd like to go back to something resembling normal

1067 00:42:52.310 --> 00:42:53.750 as I'm sure you all would,

1068 00:42:53.750 --> 00:42:55.620 but of course there's reasons to be concerned, right?

1069 00:42:55.620 --> 00:42:58.340 If efficacy is declared early in particular,

1070 00:42:58.340 --> 00:43:00.630 if that means that blinded follow-up

1071 00:43:00.630 --> 00:43:01.840 in a study stops, right?

1072 00:43:01.840 --> 00:43:02.910 Because that means we have no way

1073 00:43:02.910 --> 00:43:05.180 to assess how durable the vaccine is.

1074 00:43:05.180 --> 00:43:06.700 We won't be able to assess VE

1075 00:43:06.700 --> 00:43:09.490 and key subgroups that we care about.

1076 00:43:09.490 --> 00:43:11.130 We might not be able to assess VE

1077 00:43:11.130 --> 00:43:13.390 formally against severe end points.

1078 00:43:13.390 --> 00:43:15.260 So there's real sort of concerns

1079 00:43:15.260 --> 00:43:16.940 about stopping these trials too early,

1080 00:43:16.940 --> 00:43:18.210 and what the implications of that

1081 00:43:18.210 --> 00:43:21.138 are both for evaluating the vaccine in question,

1082 00:43:21.138 --> 00:43:23.040 but as well as how it impacts

1083 00:43:23.040 --> 00:43:25.190 the other clinical trials that are ongoing.

1084 00:43:26.120 --> 00:43:28.700 And of course in the current political climate,

1085 00:43:28.700 --> 00:43:30.710 everybody's very concerned about the role

1086 00:43:30.710 --> 00:43:33.040 political pressure might play in all of this.

1087 00:43:33.040 --> 00:43:37.466 So yeah, so it's kind of a double-edged sword in some sense

1088 00:43:37.466 --> 00:43:41.760 in terms of what happens if efficacy is declared early,

1089 00:43:41.760 --> 00:43:43.190 but I want to save just a few minutes

1090 00:43:43.190 --> 00:43:44.980 to talk about vaccine correlates 'cause I promised

1091 00:43:44.980 --> 00:43:47.170 that I would show you some math and prove to you

1092 00:43:47.170 --> 00:43:48.320 that I'm a real statistician.

1093 00:43:48.320 --> 00:43:50.800 So let's do a little bit of that.

1094 00:43:50.800 --> 00:43:52.520 So again, we're kind of shifting gears here.

1095 00:43:52.520 --> 00:43:54.650 So that's the end of sort of talking about the primary

1096 00:43:54.650 --> 00:43:56.290 analysis of these trials,

1097 00:43:56.290 --> 00:43:58.380 what's gonna lead to their licensure.

1098 00:43:58.380 --> 00:44:00.190 And the correlates of protection

1099 00:44:00.190 --> 00:44:02.300 is sort of a key secondary analysis

1100 00:44:02.300 --> 00:44:04.220 and so why is it so important

1101 00:44:04.220 --> 00:44:07.330 that we're able to establish correlates of protection?

1102 00:44:07.330 --> 00:44:08.510 Well, because it's gonna speed up

1103 00:44:08.510 --> 00:44:11.640 the whole vaccine development process.

1104 00:44:11.640 --> 00:44:14.210 So again, a correlative protection is really just,

1105 00:44:14.210 --> 00:44:17.530 it's an immune response and really an assay

1106 00:44:17.530 --> 00:44:20.150 to measure that immune response that's been validated

1107 00:44:20.150 --> 00:44:22.710 to reliably predict vaccine efficacy.

1108 00:44:22.710 --> 00:44:25.130 So why is that so important?

1109 00:44:25.130 --> 00:44:27.750 Well, basically what we're hoping to achieve

1110 00:44:27.750 --> 00:44:29.240 is the establishment of a surrogate

1111 00:44:29.240 --> 00:44:32.020 endpoint for COVID disease right?

1112 00:44:32.020 --> 00:44:34.350 So I've sort of mentioned the numbers that we're talking

1113 00:44:34.350 --> 00:44:36.120 about in these phase three trials,

1114 00:44:36.120 --> 00:44:39.640 enrolling 30,000 participants, 60,000 participants

1115 00:44:39.640 --> 00:44:41.743 and ending up with one or two years of followup, right.

1116 00:44:41.743 --> 00:44:44.130 Just to be able to answer the primary question, right.

1117 00:44:44.130 --> 00:44:47.730 Does the vaccine prevent infection and/or disease?

1118 00:44:47.730 --> 00:44:50.070 So that's a huge, expensive clinical trial.

1119 00:44:50.070 --> 00:44:52.320 It takes a long time to get an answer

1120 00:44:52.320 --> 00:44:56.080 and so it would be very nice if all we had to do right

1121 00:44:56.080 --> 00:44:58.960 was give people the doses of vaccine that they need,

1122 00:44:58.960 --> 00:45:02.180 wait two weeks and measure their immune response

1123 00:45:02.180 --> 00:45:04.540 and understand does that vaccine work or not.

1124 00:45:04.540 --> 00:45:07.385 That would be a much faster vaccine development process

1125 00:45:07.385 --> 00:45:08.900 than where we're currently at

1126 00:45:08.900 --> 00:45:11.130 in having to run these enormous phase three trials.

1127 00:45:11.130 --> 00:45:14.460 So it's valuable for establishing a surrogate endpoint.

1128 00:45:14.460 --> 00:45:17.480 It's also valuable for accelerating approval

1129 00:45:17.480 --> 00:45:21.810 of vaccines that have been licensed in certain populations,

1130 00:45:21.810 --> 00:45:22.643 but not others.

1131 00:45:22.643 --> 00:45:25.100 For example, I mentioned that these phase three trials

1132 00:45:25.100 --> 00:45:26.720 are mostly being conducted in adults.

1133 00:45:26.720 --> 00:45:30.000 Well, what if we want to also obtain licensure for use

1134 00:45:30.000 --> 00:45:31.870 of this vaccine in children?

1135 00:45:31.870 --> 00:45:34.260 Well, if we had an established immune correlate

1136 00:45:34.260 --> 00:45:35.093 we wouldn't have to do

1137 00:45:35.093 --> 00:45:37.050 a large randomized trial in children.

1138 00:45:37.050 --> 00:45:39.290 We could do it just a small immunogenicity study

1139 00:45:39.290 --> 00:45:42.137 and use the correlates results to bridge the VE

1140 00:45:42.137 --> 00:45:44.587 that we observed from the phase three trial.

1141 00:45:44.587 --> 00:45:47.280 That's the immune response that we've observed

1142 00:45:47.280 --> 00:45:49.260 in these children or pregnant women for example are

1143 00:45:49.260 --> 00:45:51.210 another key population they're being

1144 00:45:51.210 --> 00:45:53.420 excluded from these phase three trials

1145 00:45:53.420 --> 00:45:55.440 but we'd like to understand if these vaccines

1146 00:45:55.440 --> 00:45:58.650 are safe and effective in those women as well.

1147 00:45:58.650 --> 00:46:01.130 So really this is one of the key goals

1148 00:46:01.130 --> 00:46:04.900 of this whole OWS program and the key role

1149 00:46:04.900 --> 00:46:07.470 that we're playing in the CoVPN is developing

1150 00:46:08.322 --> 00:46:11.940 the sampling plan and the statistical analysis plan

1151 00:46:11.940 --> 00:46:14.090 for the immune correlate studies

1152 00:46:14.090 --> 00:46:16.620 and so it's just a little bit of the statistical issues

1153 00:46:16.620 --> 00:46:20.140 that we're dealing with in these trials, right,

1154 00:46:20.140 --> 00:46:22.140 is that sort of running assays

1155 00:46:22.140 --> 00:46:26.210 so running these immuno assays on 30,000, 60,000 individuals

1156 00:46:26.210 --> 00:46:29.900 takes a long time, it's expensive, and as it turns out,

1157 00:46:29.900 --> 00:46:33.820 it's really overkill in terms of statistical power.

1158 00:46:33.820 --> 00:46:35.310 So we can actually be a little bit more

1159 00:46:35.310 --> 00:46:39.990 clever about how we design these correlate studies in order

1160 00:46:39.990 --> 00:46:41.850 to get answers faster and more cheaply.

1161 00:46:41.850 --> 00:46:45.070 So the way we do this is we use a case cohort design.

1162 00:46:45.070 --> 00:46:46.690 So we're not gonna measure immune responses

1163 00:46:46.690 --> 00:46:47.860 in all trial participants,

1164 00:46:47.860 --> 00:46:49.800 we're gonna measure them in a sub cohort

1165 00:46:49.800 --> 00:46:51.280 and that sub cohort will consist

1166 00:46:51.280 --> 00:46:53.720 of a stratified random sub cohort.

1167 00:46:53.720 --> 00:46:56.100 So we're gonna be sampling individuals randomly

1168 00:46:56.100 --> 00:46:58.370 based on their baseline infection status.

1169 00:46:58.370 --> 00:47:01.140 Were you infected with SARS-CoV-2 in the past?

1170 00:47:01.140 --> 00:47:05.387 Based on your race, ethnicity, and based on age.

1171 00:47:06.780 --> 00:47:08.760 And so based on that, we'll take a random draw

1172 00:47:08.760 --> 00:47:12.133 of the trial population, about 1600 individuals,

1173 00:47:13.160 --> 00:47:18.160 excuse me and everyone so I should mention right

1174 00:47:18.930 --> 00:47:22.000 in the trial design everybody is having their blood drawn.

1175 00:47:22.000 --> 00:47:23.690 And right now we're talking about whose blood

1176 00:47:23.690 --> 00:47:26.500 are we gonna use to measure these immune responses?

1177 00:47:26.500 --> 00:47:28.930 So we're gonna measure it in a random sample

1178 00:47:28.930 --> 00:47:31.410 and then we're gonna wait until the trial is over

1179 00:47:31.410 --> 00:47:34.870 or until one of these interim analysis concludes efficacy

1180 00:47:34.870 --> 00:47:37.910 and we're gonna measure immune responses

1181 00:47:37.910 --> 00:47:39.240 in all of the end points, right?

1182 00:47:39.240 --> 00:47:41.890 Remember that like power in these analyses is drive

1183 00:47:41.890 --> 00:47:45.620 by the individuals in which we observe endpoints.

1184 00:47:45.620 --> 00:47:47.180 So we're gonna make sure we get immune responses

1185 00:47:47.180 --> 00:47:49.310 in all the end point data, as in addition

1186 00:47:49.310 --> 00:47:52.690 to this random sub cohort and it turns out that that's about

1187 00:47:52.690 --> 00:47:56.920 as statistically efficient as running the immune assays

1188 00:47:56.920 --> 00:47:58.890 on all 30,000 individuals in the trial.

1189 00:47:58.890 --> 00:48:01.120 So this is this kind of classic case cohort design

1190 00:48:01.120 --> 00:48:04.000 that Ross Prentice has been writing about for years

1191 00:48:04.000 --> 00:48:06.160 that Norman Breslow did some sort of pioneering work

1192 00:48:06.160 --> 00:48:10.510 on in the 2000s and I'll just talk a little bit about sort

1193 00:48:10.510 --> 00:48:13.280 of how this complicates our life as statisticians

1194 00:48:13.280 --> 00:48:16.040 and then maybe we'll leave a few minutes for questions.

1195 00:48:16.040 --> 00:48:17.610 So here's the math, we made it.

1196 00:48:17.610 --> 00:48:19.530 Well, the moment you've all been waiting for it

1197 00:48:19.530 --> 00:48:20.880 to see some math.

1198 00:48:20.880 --> 00:48:23.070 So just introducing, you know,

1199 00:48:23.070 --> 00:48:26.000 why is this sampling design challenging

1200 00:48:26.000 --> 00:48:28.740 from a perspective of generating estimators, right?

1201 00:48:28.740 --> 00:48:31.160 Well, we can sort of immediately see that this isn't

1202 00:48:31.160 --> 00:48:34.790 a totally random sample of the trial population, right?

1203 00:48:34.790 --> 00:48:38.290 In particular we've over-sampled the individuals who end up

1204 00:48:38.290 --> 00:48:42.150 getting diseased and it's fairly obvious

1205 00:48:42.150 --> 00:48:44.900 that those individuals have potential to be very different

1206 00:48:44.900 --> 00:48:47.200 than a randomly selected individual in the population.

1207 00:48:47.200 --> 00:48:48.620 So we have a bias sub sample.

1208 00:48:48.620 --> 00:48:51.950 So we need some statistical methodology to try to back out,

1209 00:48:51.950 --> 00:48:53.520 you know, whatever this parameter is.

1210 00:48:53.520 --> 00:48:56.150 We want to be estimating it in the whole trial population,

1211 00:48:56.150 --> 00:48:58.560 not just in this biased sub samples.

1212 00:48:58.560 --> 00:49:00.200 So how do we do that?

1213 00:49:00.200 --> 00:49:02.100 So just a quick notation here,

1214 00:49:02.100 --> 00:49:04.140 let's call W baseline covariates,

1215 00:49:04.140 --> 00:49:06.830 A is a binary vaccine assignment,

1216 00:49:06.830 --> 00:49:11.340 Y is your binary COVID endpoint for example

1217 00:49:11.340 --> 00:49:13.360 and then we'll introduce this sort of indicators.

1218 00:49:13.360 --> 00:49:17.570 Delta is one, if you're selected into this immune response

1219 00:49:17.570 --> 00:49:20.050 sub cohort, either because you were a case,

1220 00:49:20.050 --> 00:49:23.190 you were an end point or because you were randomly selected

1221 00:49:23.190 --> 00:49:24.600 into the cohort.

1222 00:49:24.600 --> 00:49:27.053 And then we'll call S your immune response.

1223 00:49:27.890 --> 00:49:30.870 And then we'll just say, we'll represent this as Delta S,

1224 00:49:30.870 --> 00:49:32.800 which just means we'll arbitrarily set everybody

1225 00:49:32.800 --> 00:49:35.930 who's not in our sub cohorts immune response to be zero,

1226 00:49:35.930 --> 00:49:38.170 that's arbitrary doesn't really matter.

1227 00:49:38.170 --> 00:49:40.530 So let's talk about how estimation would happen.

1228 00:49:40.530 --> 00:49:42.980 So let's pick a very simple parameter, right?

1229 00:49:42.980 --> 00:49:45.110 Let's just say that we want to know what's the overall

1230 00:49:45.110 --> 00:49:47.230 immune response in the whole population,

1231 00:49:47.230 --> 00:49:49.630 not a particularly interesting parameter

1232 00:49:49.630 --> 00:49:51.200 for actually measuring correlates,

1233 00:49:51.200 --> 00:49:53.660 but just to motivate the types of statistical approaches

1234 00:49:53.660 --> 00:49:56.090 that we use in these settings.

1235 00:49:56.090 --> 00:49:58.920 So how can we control for the bias of the sampling design?

1236 00:49:58.920 --> 00:50:00.730 Well, one of the most straightforward ways

1237 00:50:00.730 --> 00:50:02.050 is to use the tried and true

1238 00:50:02.050 --> 00:50:04.810 Horvitz-Thompson or IPTW estimator, right.

1239 00:50:04.810 --> 00:50:07.830 Where we're just taking basically a sample mean

1240 00:50:07.830 --> 00:50:11.200 but all our observations are sort of inverse weighted

1241 00:50:11.200 --> 00:50:15.920 by their probability of being sampled into this sub cohort.

1242 00:50:15.920 --> 00:50:17.997 And so that's, IPTW estimator if you're in causal inference,

1243 00:50:17.997 --> 00:50:19.460 you're very familiar with this.

1244 00:50:19.460 --> 00:50:20.710 If you're in survey sampling,

1245 00:50:20.710 --> 00:50:22.030 very familiar with this.

1246 00:50:22.030 --> 00:50:26.850 Very classical way of adjusting for this selection bias.

1247 00:50:26.850 --> 00:50:28.000 It turns out that there's ways

1248 00:50:28.000 --> 00:50:29.830 that we can be more efficient in doing this.

1249 00:50:29.830 --> 00:50:33.086 We can use augmented estimators, AIPTW estimators.

1250 00:50:33.086 --> 00:50:36.683 And the key idea there is that we take the IPTW estimator

1251 00:50:36.683 --> 00:50:39.180 and we add a little bit of something to it

1252 00:50:39.180 --> 00:50:40.850 and the key thing is that that little bit

1253 00:50:40.850 --> 00:50:45.850 of something involves a regression of S the immune response

1254 00:50:45.920 --> 00:50:49.970 onto the covariates that were used to sample individuals

1255 00:50:49.970 --> 00:50:52.140 into the sub cohort.

1256 00:50:52.140 --> 00:50:53.670 And so what's the intuition as

1257 00:50:53.670 --> 00:50:55.700 to why this is more efficient?

1258 00:50:55.700 --> 00:50:59.050 Well, you can imagine what if we had a perfect predictor

1259 00:50:59.050 --> 00:51:01.160 of S measured at baseline, right?

1260 00:51:01.160 --> 00:51:04.940 Then this regression here is essentially imputing

1261 00:51:04.940 --> 00:51:06.360 the correct value of S

1262 00:51:06.360 --> 00:51:08.860 for every single person in the population.

1263 00:51:08.860 --> 00:51:11.480 So it's kind of like we're getting more data

1264 00:51:11.480 --> 00:51:14.710 in some sense, and the nice thing about

1265 00:51:14.710 --> 00:51:16.580 these approaches, these AIPTW approaches

1266 00:51:16.580 --> 00:51:18.440 is that they're double robust and so again,

1267 00:51:18.440 --> 00:51:21.100 if you work in causal inference a very familiar idea,

1268 00:51:21.100 --> 00:51:22.630 and it turns out because we know

1269 00:51:22.630 --> 00:51:25.000 the sampling probability by design,

1270 00:51:25.000 --> 00:51:28.230 this regression doesn't have to be consistently estimated

1271 00:51:28.230 --> 00:51:29.840 in order to obtain a consistent estimate

1272 00:51:29.840 --> 00:51:30.730 of the parameter measures.

1273 00:51:30.730 --> 00:51:32.960 So it's this really nice sort of double robustness property

1274 00:51:32.960 --> 00:51:34.930 that says, yeah, you might be turned off

1275 00:51:34.930 --> 00:51:36.110 from this augmented estimator

1276 00:51:36.110 --> 00:51:37.740 because you have to do a little bit of extra work,

1277 00:51:37.740 --> 00:51:40.300 you have to fit a regression model say,

1278 00:51:40.300 --> 00:51:41.900 and maybe you're worried about missspecifying

1279 00:51:41.900 --> 00:51:44.220 that regression well it turns out that because the sampling

1280 00:51:44.220 --> 00:51:45.800 probabilities are known by design,

1281 00:51:45.800 --> 00:51:47.430 you don't have to concern yourself with that.

1282 00:51:47.430 --> 00:51:50.450 So it turns out you can use any old regression estimator

1283 00:51:50.450 --> 00:51:52.540 here and still end up with a consistent estimate

1284 00:51:52.540 --> 00:51:54.240 of the parameter of interest.

1285 00:51:54.240 --> 00:51:55.290 And so we're applying this

1286 00:51:55.290 --> 00:51:57.100 to much more interesting parameters.

1287 00:51:57.100 --> 00:51:58.520 So we had a paper come out recently

1288 00:51:58.520 --> 00:52:00.920 in biometrics that's linked here

1289 00:52:00.920 --> 00:52:03.210 where we're starting to study a sort of causal inference

1290 00:52:03.210 --> 00:52:05.650 flavored parameters in this context,

1291 00:52:05.650 --> 00:52:07.770 things that we can really use to pin down,

1292 00:52:07.770 --> 00:52:10.082 you know, mechanisms of these vaccines working.

1293 00:52:10.082 --> 00:52:12.626 So, in this case, we're studying sort of the effect

1294 00:52:12.626 --> 00:52:16.010 of a stochastic intervention, we call it.

1295 00:52:16.010 --> 00:52:17.980 So it's sort of saying what would happen

1296 00:52:17.980 --> 00:52:19.830 if we took everybody's immune response,

1297 00:52:19.830 --> 00:52:22.420 this particular immune response that we observed,

1298 00:52:22.420 --> 00:52:24.520 and we shifted it up just a little bit

1299 00:52:24.520 --> 00:52:26.470 or we shifted it down just a little bit.

1300 00:52:26.470 --> 00:52:29.580 How would that impact the risk of disease amongst

1301 00:52:29.580 --> 00:52:30.413 the vaccinated individuals?

1302 00:52:30.413 --> 00:52:33.770 So that's what this big, gnarly parameter is right here.

1303 00:52:33.770 --> 00:52:35.240 And so you ended up looking at a plot

1304 00:52:35.240 --> 00:52:36.110 that's kind of like this.

1305 00:52:36.110 --> 00:52:38.540 So this is from an HIV vaccine trial.

1306 00:52:38.540 --> 00:52:41.770 So at zero we're saying that's just the observed risk

1307 00:52:41.770 --> 00:52:44.160 of the trial and as we move left we're saying,

1308 00:52:44.160 --> 00:52:47.307 what would the risk be if we decreased your immune response?

1309 00:52:47.307 --> 00:52:48.810 And so we can see in this example,

1310 00:52:48.810 --> 00:52:51.770 we found that the risk would be increasing, right.

1311 00:52:51.770 --> 00:52:53.360 And then if we're moving to the right

1312 00:52:53.360 --> 00:52:56.587 is what would happen if we increase your immune response.

1313 00:52:56.587 --> 00:52:58.530 And so we're kind of getting at something

1314 00:52:58.530 --> 00:53:02.670 that's like a controlled effects mediation type parameter

1315 00:53:02.670 --> 00:53:06.210 with this approach and so we're working out some

1316 00:53:06.210 --> 00:53:10.370 of the details of the correlates plan currently

1317 00:53:10.370 --> 00:53:11.460 and so when that's done

1318 00:53:11.460 --> 00:53:13.070 we'll have it available for public comment.

1319 00:53:13.070 --> 00:53:14.490 And again, we're academics, right?

1320 00:53:14.490 --> 00:53:16.350 So we'll do it all open science.

1321 00:53:16.350 --> 00:53:18.270 And then I'll just say like two words of conclusion

1322 00:53:18.270 --> 00:53:20.690 and I'll shut up and leave some time for questions.

1323 00:53:20.690 --> 00:53:22.970 So there's been a big concern

1324 00:53:22.970 --> 00:53:26.000 in the current political climate that we're gonna sneak

1325 00:53:26.000 --> 00:53:28.113 something through, that something's gonna be approved

1326 00:53:28.113 --> 00:53:32.018 without sort of the standard amount of evidence

1327 00:53:32.018 --> 00:53:33.113 that would be required, right.

1328 00:53:33.113 --> 00:53:36.010 That there's political interference at the FDA

1329 00:53:36.010 --> 00:53:38.560 and from where I sit, you know,

1330 00:53:38.560 --> 00:53:40.520 I can say that the science behind the vaccine

1331 00:53:40.520 --> 00:53:43.059 development program for COVID is extremely rigorous.

1332 00:53:43.059 --> 00:53:45.900 These are exactly the type of people who you would want

1333 00:53:45.900 --> 00:53:47.969 in charge of this decision making process

1334 00:53:47.969 --> 00:53:51.270 and the type of people that will raise red flags

1335 00:53:51.270 --> 00:53:54.190 as soon as sort of the process goes off the rails.

1336 00:53:54.190 --> 00:53:56.870 So right now I feel good about where things stand.

1337 00:53:56.870 --> 00:54:00.376 Of course, I watch presidential debates and hear, you know,

1338 00:54:00.376 --> 00:54:03.620 garbage science coming out and I get a little bit concerned,

1339 00:54:03.620 --> 00:54:05.010 but from where I sit right now,

1340 00:54:05.010 --> 00:54:07.040 everything's looking pretty good.

1341 00:54:07.040 --> 00:54:09.075 So overall, I'd say that the increased transparency

1342 00:54:09.075 --> 00:54:10.960 by releasing these protocols

1343 00:54:10.960 --> 00:54:13.490 has been good for scientists and consumers.

1344 00:54:13.490 --> 00:54:15.110 We want to bring vaccines to market,

1345 00:54:15.110 --> 00:54:16.910 but we also want people to trust those vaccine

1346 00:54:16.910 --> 00:54:20.480 so increasing transparency in whatever way we can is great.

1347 00:54:20.480 --> 00:54:23.050 And then finally, the final point is that a lot of these

1348 00:54:23.050 --> 00:54:24.040 issues that I've talked about,

1349 00:54:24.040 --> 00:54:25.580 how do we do interim monitoring, right?

1350 00:54:25.580 --> 00:54:28.030 What's the right end point to be studying?

1351 00:54:28.030 --> 00:54:29.590 What's the right S demand, right?

1352 00:54:29.590 --> 00:54:31.580 These are really hard decisions

1353 00:54:31.580 --> 00:54:34.230 and there are no right answers.

1354 00:54:34.230 --> 00:54:36.650 And so one of the things that's been a little bit

1355 00:54:36.650 --> 00:54:39.610 disconcerting or disheartening to me

1356 00:54:39.610 --> 00:54:42.530 is the extent to which in the pandemic era,

1357 00:54:42.530 --> 00:54:45.860 academic debates have been made very much public

1358 00:54:45.860 --> 00:54:49.020 and I'm not against academic debates.

1359 00:54:49.020 --> 00:54:52.080 It's just that most individuals aren't used to seeing them.

1360 00:54:52.080 --> 00:54:55.330 And so what I'm worried is happening is that people

1361 00:54:55.330 --> 00:54:59.760 see high profile academics debating these challenging

1362 00:54:59.760 --> 00:55:01.870 problems where there's no real right answer.

1363 00:55:01.870 --> 00:55:03.390 And they're saying, well, these guys don't know

1364 00:55:03.390 --> 00:55:04.850 what they're talking about.

1365 00:55:04.850 --> 00:55:07.810 So I think as academics and public health professionals

1366 00:55:07.810 --> 00:55:09.920 in this pandemic, one thing that we can do

1367 00:55:09.920 --> 00:55:12.344 is just to be very careful in how we're presenting,

1368 00:55:12.344 --> 00:55:15.060 you know, the science that we're doing

1369 00:55:15.060 --> 00:55:16.890 and acknowledge when there's not a right answer,

1370 00:55:16.890 --> 00:55:18.590 that you're presenting your opinion.

1371 00:55:18.590 --> 00:55:20.850 And that there is some validity, right?

1372 00:55:20.850 --> 00:55:23.420 That this is very gray, unfortunately,

1373 00:55:23.420 --> 00:55:25.260 that there's nothing black and white here.

1374 00:55:25.260 --> 00:55:27.640 So maybe that's a controversial statement to end on,

1375 00:55:27.640 --> 00:55:29.940 but I'll end there and then thanks again to Fan

1376 00:55:29.940 --> 00:55:31.527 for giving me the opportunity to talk

1377 00:55:31.527 --> 00:55:34.840 and I'm happy to take questions as there's time.

1378 00:55:34.840 --> 00:55:36.430 I don't have anything scheduled after this,

1379 00:55:36.430 --> 00:55:39.200 so I can stay a few minutes over as would be helpful.

1380 00:55:39.200 --> 00:55:40.183 So thanks again.

1381 00:55:41.386 --> 00:55:43.730 - [Fan] Thank you David for this very nice talk.

1382 00:55:43.730 --> 00:55:46.460 I think we do have three to four minutes for questions

1383 00:55:47.383 --> 00:55:50.423 from the audience, if there's any.

1384 00:55:53.890 --> 00:55:55.390 - [Woman] Hi David, I have a question

1385 00:55:55.390 --> 00:56:00.100 'cause right now for COVID situation and because of the time

1386 00:56:00.100 --> 00:56:03.720 and the faster progress of the disease

1387 00:56:03.720 --> 00:56:07.870 and it's a hard to keep the standard method,

1388 00:56:07.870 --> 00:56:12.870 but do you have other proofed vaccine for other disease

1389 00:56:13.890 --> 00:56:18.890 and have a quick trial have a similar way as COVID

1390 00:56:19.450 --> 00:56:23.590 and apply the method you're using right now

1391 00:56:23.590 --> 00:56:26.910 and we have standard results already

1392 00:56:26.910 --> 00:56:31.840 and then compare to see how good the current method is.

1393 00:56:31.840 --> 00:56:33.533 So that's my question.

1394 00:56:35.110 --> 00:56:37.217 - [David] Yeah it's an interesting question.

1395 00:56:37.217 --> 00:56:39.320 So let me try to restate, so you're saying,

1396 00:56:39.320 --> 00:56:42.090 are there any lessons from vaccine development

1397 00:56:42.090 --> 00:56:44.136 that we can try to draw from here

1398 00:56:44.136 --> 00:56:47.590 to evaluate our methodology, whether it work?

1399 00:56:47.590 --> 00:56:50.933 - [Woman] Yes, from other vaccines.

1400 00:56:51.840 --> 00:56:54.920 - [David] So I guess what I would say is that at this stage,

1401 00:56:54.920 --> 00:56:58.280 in phase three vaccines, these phase three trials

1402 00:56:58.280 --> 00:56:59.820 look completely normal.

1403 00:56:59.820 --> 00:57:02.680 So I would say the process of getting to the phase three

1404 00:57:02.680 --> 00:57:04.920 looked very different and much more accelerated

1405 00:57:04.920 --> 00:57:07.170 in terms of kind of squashing together

1406 00:57:07.170 --> 00:57:11.300 phase one and phase two in terms of the manufacturing,

1407 00:57:11.300 --> 00:57:13.410 but in terms of what's happening in a phase three trial,

1408 00:57:13.410 --> 00:57:14.760 this is probably the phase three trial

1409 00:57:14.760 --> 00:57:18.220 that would be done outside of the setting of a pandemic.

1410 00:57:18.220 --> 00:57:20.220 Maybe the interim analysis would be a little bit

1411 00:57:20.220 --> 00:57:23.390 less aggressive for some of these companies, but really,

1412 00:57:23.390 --> 00:57:26.614 I think the approaches that the companies are taking

1413 00:57:26.614 --> 00:57:30.903 would be fairly standard even in any other vaccine context.

1414 00:57:34.831 --> 00:57:36.842 - -Woman] Yeah. I mean, even though

1415 00:57:36.842 --> 00:57:39.913 for the established vaccine,

1416 00:57:40.820 --> 00:57:43.260 there could be some field trial

1417 00:57:43.260 --> 00:57:47.150 and that they also went through a phase three,

1418 00:57:47.150 --> 00:57:50.540 but you can do the similar thing to enhance,

1419 00:57:50.540 --> 00:57:55.000 to see whether it is possible to pass the current protocol

1420 00:57:56.512 --> 00:57:59.473 and become some sort of false positive.

1421 00:58:02.350 --> 00:58:05.543 - [David] Yeah and, you know, I think speaking,

1422 00:58:07.930 --> 00:58:09.320 I mean, speaking of failed vaccines,

1423 00:58:09.320 --> 00:58:11.497 as someone who works in HIV vaccines,

1424 00:58:11.497 --> 00:58:14.700 we're very familiar with failure and learning from that.

1425 00:58:14.700 --> 00:58:17.497 So again, I think the people who are running these trials

1426 00:58:17.497 --> 00:58:20.130 are sort of the right people in terms of looking out

1427 00:58:20.130 --> 00:58:22.430 for these false positive signals and so forth.

1428 00:58:24.132 --> 00:58:25.132 - [Woman] Thank you.

1429 00:58:26.513 --> 00:58:29.580 - [Fan] So I think we are just about the time

1430 00:58:29.580 --> 00:58:32.040 and I'm sure that David is happy

1431 00:58:32.040 --> 00:58:35.030 to take your questions afterwards by email.

1432 00:58:35.030 --> 00:58:37.150 So I'll thank David more time.

1433 00:58:37.150 --> 00:58:39.010 Again, thank you for sharing with us

1434 00:58:39.010 --> 00:58:41.223 and we'll see everyone again next week.

1435 00:58:43.070 --> 00:58:44.063 - [David] Thanks everybody.