WEBVTT

1 00:00:00.120 --> 00:00:00.953 <v Host>Assistant professor</v>

2 00:00:00.953 --> 00:00:02.550 in the Department of Population Health

3 00:00:02.550 --> 00:00:05.520 and in the Department of Medicine at New York University,

4 00:00:05.520 --> 00:00:08.760 Dr. Wu's research synthesizes state-of-the-art methods

5 00:00:08.760 --> 00:00:11.280 from statistics, machine learning, optimization,

6 00:00:11.280 --> 00:00:14.160 and computational science to address critical

7 00:00:14.160 --> 00:00:16.290 and far reaching issues in health services,

8 00:00:16.290 --> 00:00:18.420 research and clinical practice.

9 00:00:18.420 --> 00:00:20.070 Leveraging large scale data

10 00:00:20.070 --> 00:00:23.940 from national disease registries, administrative databases,

11 00:00:23.940 --> 00:00:27.120 electronic health records, and randomized control trials.

12 00:00:27.120 --> 00:00:29.313 Let's give a warm welcome to Dr. Wu.

13 00:00:31.472 --> 00:00:33.870 <v Dr. Wu>Thank you for the nice introduction.</v>

14 00:00:33.870 --> 00:00:38.820 And it's a great honor to be here with all of you.

15 00:00:38.820 --> 00:00:43.563 And so I'm Wenbo, I am from New York.

16 00:00:44.730 --> 00:00:49.730 I joined NYU just a bit over a year ago.

17 00:00:52.574 --> 00:00:54.840 So I think, 'cause we have so many people here,

18 00:00:54.840 --> 00:00:58.034 I think it would be good to run a promotion first.

19 00:00:58.034 --> 00:00:59.220 (Dr. Wu laughs)

20 00:00:59.220 --> 00:01:00.930 So this is our group.

21 00:01:00.930 --> 00:01:03.120 So at NYU we have,

22 00:01:03.120 --> 00:01:06.627 I mean it's a tremendously growing group

23 00:01:06.627 --> 00:01:08.490 and we have like 24 faculty

24 00:01:08.490 --> 00:01:11.970 and we're about to welcome our newest,

25 00:01:11.970 --> 00:01:16.020 like the 25th faculty member into our divisions.

26 00:01:16.020 --> 00:01:19.290 And we have 7 staff.

27 00:01:19.290 --> 00:01:21.690 We have a small PhD program,

28 00:01:21.690 --> 00:01:25.530 we have 20 PhD students and 10 postdocs.

29 00:01:25.530 --> 00:01:30.270 And we have a team of 25 research scientists.

30 00:01:30.270 --> 00:01:33.090 And part of the reason I wanna do this is

31 00:01:33.090 --> 00:01:36.210 because I wanna encourage you guys

32 00:01:36.210 --> 00:01:37.800 to apply to our PhD programs.

33 00:01:37.800 --> 00:01:39.600 So if you're interested,

34 00:01:39.600 --> 00:01:43.170 scan this QR code and you apply, okay?

35 00:01:43.170 --> 00:01:44.670 All right,

36 00:01:44.670 --> 00:01:49.670 so I have been doing things in provider profiling

37 00:01:49.980 --> 00:01:53.400 for the past for five years

38 00:01:53.400 --> 00:01:58.400 and so this is the overview of what it is.

39 00:01:58.650 --> 00:02:02.700 So provider profiling is basically the assessment,

40 00:02:02.700 --> 00:02:07.383 the evaluation of the performance of healthcare providers.

41 00:02:09.180 --> 00:02:10.200 So I listed here,

42 00:02:10.200 --> 00:02:13.143 could be say acute-care hospitals.

43 00:02:14.454 --> 00:02:17.400 (Wu speaks indistinctly)

44 00:02:17.400 --> 00:02:20.670 This acute-care hospitals, kidney dialysis facilities,

45 00:02:20.670 --> 00:02:25.230 I have been working on other evaluations

46 00:02:25.230 --> 00:02:27.360 like organ procurement organizations,

47 00:02:27.360 --> 00:02:30.960 which is a type of organizations

48 00:02:30.960 --> 00:02:33.867 that are responsible for procuring organs

49 00:02:33.867 --> 00:02:36.390 for patients who are in great need

50 00:02:36.390 --> 00:02:38.910 of organ transplant patients.

51 00:02:38.910 --> 00:02:41.520 And the transplant centers, of course, physician, surgeons.

52 00:02:41.520 --> 00:02:43.560 So you can see,

53 00:02:43.560 --> 00:02:45.930 this includes so many different types

54 00:02:45.930 --> 00:02:49.919 of healthcare providers and stakeholders include,

55 00:02:49.919 --> 00:02:54.919 say, insurance companies, regulation, government,

56 00:02:54.930 --> 00:02:56.190 federal agencies.

57 00:02:56.190 --> 00:02:58.650 They're all interested in provider profiling,

58 00:02:58.650 --> 00:03:00.150 I will tell you why.

59 00:03:00.150 --> 00:03:05.110 Providers is basically who are doing profile evaluations

60 00:03:06.060 --> 00:03:07.830 and of course patients.

61 00:03:07.830 --> 00:03:11.490 So because they are interested in the information,

62 00:03:11.490 --> 00:03:13.240 interested in the profiling results

63 00:03:14.370 --> 00:03:17.250 so they can make care seeking decisions.

64 00:03:17.250 --> 00:03:18.090 Okay?

65 00:03:18.090 --> 00:03:22.140 And so I listed here a few outcomes,

66 00:03:22.140 --> 00:03:25.470 like emergency department encounters,

67 00:03:25.470 --> 00:03:28.500 unplanned re-hospitalizations,

68 00:03:28.500 --> 00:03:30.930 which is hospital readmissions.

69 00:03:30.930 --> 00:03:34.170 And I will jump into the details later

70 00:03:34.170 --> 00:03:36.270 and post-discharge deaths and you can,

71 00:03:36.270 --> 00:03:38.550 I mean there are so many different types of outcomes

72 00:03:38.550 --> 00:03:40.250 to consider in provider profiling.

73 00:03:41.280 --> 00:03:43.440 And one of the goals was

74 00:03:43.440 --> 00:03:46.740 to basically identify those providers

75 00:03:46.740 --> 00:03:48.600 with very bad performance in terms

76 00:03:48.600 --> 00:03:50.400 of patient-centered outcomes.

77 00:03:50.400 --> 00:03:55.110 And they can get penalization,

78 00:03:55.110 --> 00:03:58.260 like they can have payment reductions

79 00:03:58.260 --> 00:04:00.570 from government agencies.

80 00:04:00.570 --> 00:04:01.500 Okay?

81 00:04:01.500 --> 00:04:04.289 And as you can see here, this is very important.

82 00:04:04.289 --> 00:04:07.020 This is a very important business,

83 00:04:07.020 --> 00:04:10.405 and profiling can actually help

84 00:04:10.405 --> 00:04:13.140 improve evidence-based accountability

85 00:04:13.140 --> 00:04:18.090 for those providers and how facility targeted interventions

86 00:04:18.090 --> 00:04:22.083 that aimed at improving the care quality.

87 00:04:23.910 --> 00:04:24.903 Alright, so,

88 00:04:35.961 --> 00:04:36.794 so,

89 00:04:39.587 --> 00:04:42.690 this is a slide of a few example papers

90 00:04:42.690 --> 00:04:47.160 that are about evaluating hospitals across the nations.

91 00:04:47.160 --> 00:04:51.843 So they're mostly from the program called,

92 00:04:53.280 --> 00:04:55.380 Hospital Re-admission Reduction Program,

93 00:04:55.380 --> 00:04:58.500 which is a very important national level program

94 00:04:58.500 --> 00:05:00.570 that I will explain later.

95 00:05:00.570 --> 00:05:04.590 But there are just so many papers in this field.

96 00:05:04.590 --> 00:05:05.913 I mean, these are just,

97 00:05:06.953 --> 00:05:10.170 like there are publications in top,

98 00:05:10.170 --> 00:05:12.660 medical journals, analysts of internal medicine,

99 00:05:12.660 --> 00:05:14.733 and New England Journal of Medicine.

100 00:05:24.450 --> 00:05:25.283 Okay?

101 00:05:27.180 --> 00:05:29.730 So, this is another type of profiling stuff.

102 00:05:29.730 --> 00:05:31.170 So it's called physician profiling.

103 00:05:31.170 --> 00:05:36.170 Basically they wanna evaluate physicians.

104 00:05:36.180 --> 00:05:39.480 So this is, as you can see, it's a report,

105 00:05:39.480 --> 00:05:41.357 it's called the health report

106 00:05:41.357 --> 00:05:44.400 from Massachusetts Medical Society,

107 00:05:44.400 --> 00:05:45.540 which is the publisher

108 00:05:45.540 --> 00:05:46.830 of The New England Journal of Medicine.

109 00:05:46.830 --> 00:05:47.663 Okay?

110 00:05:47.663 --> 00:05:50.416 So they prepared this principles

111 00:05:50.416 --> 00:05:54.513 for profiling physician performance, I think many years ago.

112 00:05:56.130 --> 00:06:00.750 So this is a list of exemplar profiling programs

113 00:06:00.750 --> 00:06:03.870 and they are still existing.

114 00:06:03.870 --> 00:06:08.490 So the first one is an interesting state level program

115 00:06:08.490 --> 00:06:12.150 which is arguably one of the first programs.

116 00:06:12.150 --> 00:06:17.150 So it is still administered

117 00:06:18.990 --> 00:06:21.930 by the New York State of Department of Health.

118 00:06:21.930 --> 00:06:24.840 Basically they're interested in evaluating hospitals

119 00:06:24.840 --> 00:06:29.840 that do coronary artery bypass graft surgeries,

120 00:06:30.521 --> 00:06:35.310 and also PCIs and the program have been running

121 00:06:35.310 --> 00:06:38.640 for at least 20 years or so.

122 00:06:38.640 --> 00:06:41.490 And the second one is another important program,

123 00:06:41.490 --> 00:06:46.440 which was launched I think in 2003.

124 00:06:46.440 --> 00:06:47.823 And it is,

125 00:06:48.690 --> 00:06:52.380 I think it is from the one of the Federal Level Act.

126 00:06:52.380 --> 00:06:54.210 And it is currently administered

127 00:06:54.210 --> 00:06:58.290 by the US Centers for Medicare and Medicaid Services.

128 00:06:58.290 --> 00:07:03.290 And their interest in outcomes for, again,

129 00:07:03.960 --> 00:07:08.520 30-day readmissions and mortality for a AMIs

130 00:07:08.520 --> 00:07:10.263 and the heart failure, et cetera.

131 00:07:11.400 --> 00:07:15.733 And the next one is another federal level readmission,

132 00:07:17.370 --> 00:07:18.630 federal level profiling program,

133 00:07:18.630 --> 00:07:23.610 which is also established by Affordable Care Act,

134 00:07:23.610 --> 00:07:24.840 which is Obama care.

135 00:07:24.840 --> 00:07:28.560 You guys probably know that, in 2012.

136 00:07:28.560 --> 00:07:32.520 And so, yeah, they're also interested in,

137 00:07:32.520 --> 00:07:36.990 evaluating hospitals and they will punish those hospitals

138 00:07:36.990 --> 00:07:40.320 with very bad performance in terms of payment reductions.

139 00:07:40.320 --> 00:07:41.153 Okay?

140 00:07:41.153 --> 00:07:42.930 The last one is an interesting program,

141 00:07:42.930 --> 00:07:45.810 which is kind of my focus.

142 00:07:45.810 --> 00:07:50.810 I have been working on evaluating kidney dialysis facilities

143 00:07:53.730 --> 00:07:56.220 for patients with kidney failure.

144 00:07:56.220 --> 00:08:01.220 And there are actually over 7,000 dialysis facilities

145 00:08:01.320 --> 00:08:03.570 across the nation, believe it or not.

146 00:08:03.570 --> 00:08:08.080 But this is the first to pay for performance program

147 00:08:09.240 --> 00:08:13.620 in contrast to other pay for service programs.

148 00:08:13.620 --> 00:08:14.520 Okay.

149 00:08:14.520 --> 00:08:17.100 And the program is called ESRD.

150 00:08:17.100 --> 00:08:20.850 ESRD is short for End Stage Renal Disease.

151 00:08:20.850 --> 00:08:23.580 Basically the patients with kidney failure,

152 00:08:23.580 --> 00:08:25.280 a quality incentive program, okay?

153 00:08:26.160 --> 00:08:26.993 Alright.

154 00:08:26.993 --> 00:08:30.090 So as you can see, there are so many programs,

155 00:08:30.090 --> 00:08:35.037 so many initiatives across the nation about profiling.

156 00:08:35.940 --> 00:08:40.590 And one natural question is about the,

157 00:08:40.590 --> 00:08:43.980 how the landscape of the statistical landscape

158 00:08:43.980 --> 00:08:45.870 of profiling looks like.

159 00:08:45.870 --> 00:08:49.470 And because of the importance of profiling

160 00:08:49.470 --> 00:08:51.603 and here I said,

161 00:08:52.680 --> 00:08:54.930 there are many far reaching implications

162 00:08:54.930 --> 00:08:57.687 because providers can get penalizations

163 00:08:57.687 --> 00:09:00.333 and it's high stakes.

164 00:09:01.680 --> 00:09:03.480 So it's important

165 00:09:03.480 --> 00:09:05.460 that we have principles statistical methods

166 00:09:05.460 --> 00:09:07.590 to evaluate them, right?

167 00:09:07.590 --> 00:09:10.740 So this is like two examples.

168 00:09:10.740 --> 00:09:12.531 The first,

169 00:09:12.531 --> 00:09:16.980 it's a paper published on analysts of internal medicine,

170 00:09:16.980 --> 00:09:20.463 but it is written by two statisticians.

171 00:09:21.450 --> 00:09:25.470 They are calling for the improvement

172 00:09:25.470 --> 00:09:28.410 of statistical approach in this field.

173 00:09:28.410 --> 00:09:30.450 And also the second one,

174 00:09:30.450 --> 00:09:32.340 this one is even more important

175 00:09:32.340 --> 00:09:34.950 because it is a white paper issued

176 00:09:34.950 --> 00:09:39.950 by the Committee of Presidents of Statistical Society.

177 00:09:40.110 --> 00:09:41.580 You probably know about COPS.

178 00:09:41.580 --> 00:09:46.580 So one of the most important words in the statistic field,

179 00:09:46.590 --> 00:09:49.230 it's the COPS presence of work, right?

180 00:09:49.230 --> 00:09:53.130 So this is a white paper by COPS

181 00:09:53.130 --> 00:09:58.130 and also a group of people from the CMS.

182 00:09:58.230 --> 00:10:00.090 So this is also an important work.

183 00:10:00.090 --> 00:10:01.762 It's about the statistical issues

184 00:10:01.762 --> 00:10:05.190 and assessing hospital performance.

185 00:10:05.190 --> 00:10:06.750 So as you can see,

186 00:10:06.750 --> 00:10:09.510 there are many people are interested

187 00:10:09.510 --> 00:10:13.080 in improving the statistical landscape for profiling.

188 00:10:14.430 --> 00:10:15.263 Alright,

189 00:10:15.263 --> 00:10:20.120 so this is a slight briefly introducing the existing methods

190 00:10:23.070 --> 00:10:24.810 of provider profiling.

191 00:10:24.810 --> 00:10:26.280 There are a few.

192 00:10:26.280 --> 00:10:31.110 I grouped them into like roughly four categories.

193 00:10:31.110 --> 00:10:34.410 So the first group,

194 00:10:34.410 --> 00:10:38.070 is hierarchical random-effects models,

195 00:10:38.070 --> 00:10:41.610 there are many papers in this group,

196 00:10:41.610 --> 00:10:44.970 but I just highlighted one paper in,

197 00:10:44.970 --> 00:10:48.490 I think in 1997 was published on Jassa

198 00:10:49.590 --> 00:10:53.970 by Dr. Sharon Lee Norman at Harvard Medical School.

199 00:10:53.970 --> 00:10:57.780 So it's about hierarchical random-effects models

200 00:10:57.780 --> 00:11:02.340 which is still being used in many settings.

201 00:11:02.340 --> 00:11:03.780 Especially, I mean,

202 00:11:03.780 --> 00:11:05.370 not sure whether you guys know

203 00:11:05.370 --> 00:11:08.310 that there is a group at Yale called Yale Core,

204 00:11:08.310 --> 00:11:10.593 I think Center for Outcomes Research and,

205 00:11:14.130 --> 00:11:15.720 Something. <v ->Evaluation.</v>

206 00:11:15.720 --> 00:11:17.430 <v Dr. Wu>Okay, great, thank you.</v>

207 00:11:17.430 --> 00:11:21.330 So they have been using hierarchical random-effects model

208 00:11:21.330 --> 00:11:24.120 for over 30 years, I guess.

209 00:11:24.120 --> 00:11:29.120 And the second stream of approach is fixed-effects models,

210 00:11:31.020 --> 00:11:33.663 as you can tell from the names,

211 00:11:35.670 --> 00:11:38.913 people are using like a fixed effects in the models.

212 00:11:40.260 --> 00:11:44.040 And this is one example paper,

213 00:11:44.040 --> 00:11:48.003 actually was published in 2013 by my advisors.

214 00:11:49.020 --> 00:11:53.301 And the next one is,

215 00:11:53.301 --> 00:11:56.100 I mean these groups of papers,

216 00:11:56.100 --> 00:11:59.550 they're not mutually exclusive because,

217 00:11:59.550 --> 00:12:01.470 for example, this one,

218 00:12:01.470 --> 00:12:04.140 competing risks or semi-competing risks.

219 00:12:04.140 --> 00:12:05.970 I mean there are some papers

220 00:12:05.970 --> 00:12:08.190 that use higher hierarch random-effects model

221 00:12:08.190 --> 00:12:11.580 or they're also papers using fixed-effects models.

222 00:12:11.580 --> 00:12:13.470 But they are just kind of,

223 00:12:13.470 --> 00:12:15.660 they're handling like different types of outcomes.

224 00:12:15.660 --> 00:12:18.120 So I listened here.

225 00:12:18.120 --> 00:12:20.250 And also for recurring events,

226 00:12:20.250 --> 00:12:23.460 if you take a class in survival analysis,

227 00:12:23.460 --> 00:12:25.650 you probably know that, for example,

228 00:12:25.650 --> 00:12:28.680 patient can have multiple hospitalizations in a year.

229 00:12:28.680 --> 00:12:31.230 So they are considered as recurring events.

230 00:12:31.230 --> 00:12:32.329 Okay.

231 00:12:32.329 --> 00:12:34.203 And then the last one is,

232 00:12:35.280 --> 00:12:37.290 some people are using causal inference

233 00:12:37.290 --> 00:12:42.250 and some clustering approaches to handle profiling issues.

234 00:12:43.920 --> 00:12:46.740 But these papers are relatively new,

235 00:12:46.740 --> 00:12:50.460 and this is one paper here.

236 00:12:50.460 --> 00:12:53.433 It was by all statistics, I think.

237 00:12:54.420 --> 00:12:59.130 Alright, so I wanna discuss a few limitations

238 00:12:59.130 --> 00:13:01.200 of the current landscape,

239 00:13:01.200 --> 00:13:05.070 the current statistical in profiling.

240 00:13:05.070 --> 00:13:10.070 So the first limitation is, people have been, I think,

241 00:13:10.410 --> 00:13:14.073 intensely using models with a linear predictor.

242 00:13:14.910 --> 00:13:19.468 So the limitation is this may not be true

243 00:13:19.468 --> 00:13:22.560 when we have very complex outcome

244 00:13:22.560 --> 00:13:25.080 and the factor associations.

245 00:13:25.080 --> 00:13:27.510 So this is an example.

246 00:13:27.510 --> 00:13:28.803 This figure.

247 00:13:30.480 --> 00:13:35.480 This is in my one of my papers.

248 00:13:35.700 --> 00:13:38.310 So the background,

249 00:13:38.310 --> 00:13:40.680 I'll give you a bit of background information.

250 00:13:40.680 --> 00:13:42.450 So this is about, okay,

251 00:13:42.450 --> 00:13:45.764 evaluating the effect of covid

252 00:13:45.764 --> 00:13:50.550 and the outcome is a 30 day unplanned hospital readmissions.

253 00:13:50.550 --> 00:13:53.700 So this, on the left is the surface plot.

254 00:13:53.700 --> 00:13:56.460 On the right is the conquer plot.

255 00:13:56.460 --> 00:13:58.710 As you can see,

256 00:13:58.710 --> 00:14:01.920 we are interested in the variation

257 00:14:01.920 --> 00:14:05.970 of the covid effect across, this might be too small,

258 00:14:05.970 --> 00:14:08.850 but across post discharge time,

259 00:14:08.850 --> 00:14:12.660 post discharge days and also across calendar days

260 00:14:12.660 --> 00:14:15.300 because we used data in 2020.

261 00:14:15.300 --> 00:14:20.300 So we set time zero at, I think mid-March or,

262 00:14:21.240 --> 00:14:22.440 yeah, mid-March.

263 00:14:22.440 --> 00:14:24.780 So this is April the 1st.

264 00:14:24.780 --> 00:14:29.780 And then May 1st until I think mid-October.

265 00:14:29.850 --> 00:14:34.173 So as you can see there's a lot of variation going on here.

266 00:14:35.820 --> 00:14:38.430 So the covid effect is definitely not constant here.

267 00:14:38.430 --> 00:14:43.430 So basically it means that we cannot use the linear model

268 00:14:43.590 --> 00:14:44.423 to do this.

269 00:14:44.423 --> 00:14:47.850 It's just not valid, right?

270 00:14:47.850 --> 00:14:52.850 So the second methodological limitation is existing methods

271 00:14:53.970 --> 00:14:57.360 have been historically driven by cost effective spending.

272 00:14:57.360 --> 00:14:58.193 Like,

273 00:15:00.510 --> 00:15:03.077 I think in the very first program,

274 00:15:03.077 --> 00:15:06.330 in those first early programs,

275 00:15:06.330 --> 00:15:10.440 people are interested in how to reduce costs

276 00:15:10.440 --> 00:15:13.320 by, of course they wanna improve,

277 00:15:13.320 --> 00:15:14.730 they wanna improve care quality

278 00:15:14.730 --> 00:15:19.083 but cost effectiveness is a very important factor.

279 00:15:20.130 --> 00:15:21.570 So,

280 00:15:21.570 --> 00:15:22.440 and these analysis,

281 00:15:22.440 --> 00:15:25.389 they basically combine all racial ethnic groups together

282 00:15:25.389 --> 00:15:28.173 without accounting for their heterogeneity.

283 00:15:30.540 --> 00:15:32.975 So this is an another example.

284 00:15:32.975 --> 00:15:37.170 So we basically look at the performance

285 00:15:37.170 --> 00:15:41.463 of Organ Procurement Organizations, OPOs.

286 00:15:42.360 --> 00:15:43.620 So we are interested

287 00:15:43.620 --> 00:15:48.620 in organization level transplantation rates.

288 00:15:48.900 --> 00:15:51.273 And we have data in 2020.

289 00:15:53.010 --> 00:15:54.960 So these are,

290 00:15:54.960 --> 00:15:59.400 so on the y-axis we have the normalized OPO IDs,

291 00:16:02.504 --> 00:16:06.600 and this is just like a three panels of caterpillar plots.

292 00:16:06.600 --> 00:16:11.600 And if we focus on a certain OPO, then,

293 00:16:12.180 --> 00:16:13.470 for example, in this panel,

294 00:16:13.470 --> 00:16:16.200 this is a panel for white patients.

295 00:16:16.200 --> 00:16:19.050 And if you look at this is,

296 00:16:19.050 --> 00:16:20.730 I know this is a little bit small,

297 00:16:20.730 --> 00:16:23.490 but this is OPO 30 and this,

298 00:16:23.490 --> 00:16:27.180 the conference interval is above the national rate

299 00:16:27.180 --> 00:16:28.620 for white patients.

300 00:16:28.620 --> 00:16:32.220 So it's significantly better than the national average.

301 00:16:32.220 --> 00:16:36.900 But if you look at the this panel,

302 00:16:36.900 --> 00:16:39.510 this is also OPO 30

303 00:16:39.510 --> 00:16:43.800 and we have the confidence interval being lower

304 00:16:43.800 --> 00:16:46.380 than the national average for black patients.

305 00:16:46.380 --> 00:16:51.380 And this is a panel for Asian Americans

306 00:16:51.690 --> 00:16:52.950 and Pacific Islanders.

307 00:16:52.950 --> 00:16:57.510 We also have the same issue going on here for OPO 30.

308 00:16:57.510 --> 00:17:02.510 So as you can see, there's definitely racial disparity here,

309 00:17:03.630 --> 00:17:08.630 but this was never examined in those early programs.

310 00:17:10.590 --> 00:17:14.610 So this is an limitation of course.

311 00:17:14.610 --> 00:17:16.263 And the last one is,

312 00:17:17.220 --> 00:17:19.890 there is a lack of a unifying framework

313 00:17:19.890 --> 00:17:24.049 to accommodate different provider profiling objectives

314 00:17:24.049 --> 00:17:27.480 and the different performance benchmarks.

315 00:17:27.480 --> 00:17:31.350 I will give you like four different examples.

316 00:17:31.350 --> 00:17:32.183 The first one,

317 00:17:33.690 --> 00:17:36.810 I tried to make the notation very easy.

318 00:17:36.810 --> 00:17:41.810 So say we have a random-effects model here.

319 00:17:42.330 --> 00:17:45.120 We just consider a binary outcome.

320 00:17:45.120 --> 00:17:46.860 Y can be zero or one.

321 00:17:46.860 --> 00:17:47.693 Okay?

322 00:17:47.693 --> 00:17:51.660 And we basically use the logistic regression, here.

323 00:17:51.660 --> 00:17:56.190 So this gamma i, it's a sum of two things.

324 00:17:56.190 --> 00:17:58.230 The first one is mu as the mean effect.

325 00:17:58.230 --> 00:18:02.583 And the second one is ID normally distributed,

326 00:18:05.071 --> 00:18:06.510 a random variable, okay?

327 00:18:06.510 --> 00:18:08.400 And we can construct a type of,

328 00:18:08.400 --> 00:18:10.410 we call it standardized measure.

329 00:18:10.410 --> 00:18:12.990 It's Oi divided by Ei,

330 00:18:12.990 --> 00:18:16.860 O is just a sum of all those YIJs.

331 00:18:16.860 --> 00:18:19.230 And the Ei is the,

332 00:18:19.230 --> 00:18:22.770 basically the sig y function transformation

333 00:18:22.770 --> 00:18:25.080 of mu plus beta.

334 00:18:25.080 --> 00:18:26.610 Okay?

335 00:18:26.610 --> 00:18:29.700 So here, if you look at the model,

336 00:18:29.700 --> 00:18:31.200 we have gamma I here,

337 00:18:31.200 --> 00:18:35.490 but when we calculate the expected number of events

338 00:18:35.490 --> 00:18:39.033 or outcomes, we replace this with the mean.

339 00:18:40.080 --> 00:18:40.913 Okay?

340 00:18:40.913 --> 00:18:43.740 So this is the first example

341 00:18:43.740 --> 00:18:45.690 of course using random effects models.

342 00:18:45.690 --> 00:18:49.410 But if we look at the fixed effects model,

343 00:18:49.410 --> 00:18:51.870 we have the similar formulation here,

344 00:18:51.870 --> 00:18:53.880 but here because this is a fixed-effects model,

345 00:18:53.880 --> 00:18:57.690 gamma I is just unknown fixed effect, okay?

346 00:18:57.690 --> 00:19:01.530 And if we define gamma,

347 00:19:01.530 --> 00:19:05.100 start to be the median of gamma, this is a vector actually.

348 00:19:05.100 --> 00:19:08.160 So it's a vector of vault fixed-effects.

349 00:19:08.160 --> 00:19:12.480 Then this is basically the median of vault provider effects

350 00:19:12.480 --> 00:19:13.710 or fixed effects.

351 00:19:13.710 --> 00:19:16.980 And so we can also construct this standardized measure,

352 00:19:16.980 --> 00:19:21.980 but this time, this E is defined as this,

353 00:19:22.410 --> 00:19:26.410 and this is gamma star.

354 00:19:26.410 --> 00:19:30.240 So we basically use the median of all fixed effects

355 00:19:30.240 --> 00:19:32.670 to construct the standardized measure.

356 00:19:32.670 --> 00:19:33.503 Okay?

357 00:19:33.503 --> 00:19:35.850 So now we have two cases.

358 00:19:35.850 --> 00:19:39.300 One is, okay, we use the, oops,

359 00:19:39.300 --> 00:19:44.300 we use mu, which is the mean of all provider effects,

360 00:19:44.430 --> 00:19:46.470 although it's a random effects model.

361 00:19:46.470 --> 00:19:48.069 And,

362 00:19:48.069 --> 00:19:53.069 here we have median of all fixed provider effects, okay?

363 00:19:53.520 --> 00:19:55.230 So these are two cases,

364 00:19:55.230 --> 00:19:58.380 basically two types of models that have been used before.

365 00:19:58.380 --> 00:20:03.380 And next one is, and some causal papers,

366 00:20:04.020 --> 00:20:08.610 they can use a selected set of provider,

367 00:20:08.610 --> 00:20:10.637 it could be a single provider,

368 00:20:10.637 --> 00:20:13.980 let's say, I'm a a hospital administrator,

369 00:20:13.980 --> 00:20:15.360 I wanna see, okay,

370 00:20:15.360 --> 00:20:19.050 whether my hospital is performing better or worse

371 00:20:19.050 --> 00:20:21.270 than another hospital,

372 00:20:21.270 --> 00:20:25.050 then of course I can use my hospital as the benchmark,

373 00:20:25.050 --> 00:20:28.560 as the reference and compare all other hospital

374 00:20:28.560 --> 00:20:30.090 with my hospital, okay?

375 00:20:30.090 --> 00:20:31.950 So this is the first case.

376 00:20:31.950 --> 00:20:35.880 We can just choose a single hospital or provider

377 00:20:35.880 --> 00:20:37.230 as the benchmark.

378 00:20:37.230 --> 00:20:41.910 And the second case is we can group a few providers,

379 00:20:41.910 --> 00:20:45.480 hospitals in the specific geographic region together

380 00:20:45.480 --> 00:20:48.750 and to form a benchmark, this is also doable, okay?

381 00:20:48.750 --> 00:20:53.280 And it is actually used in the paper.

382 00:20:53.280 --> 00:20:57.600 The last one is, we can basically treat all hospitals,

383 00:20:57.600 --> 00:20:59.760 you can group all hospitals together

384 00:20:59.760 --> 00:21:02.070 into a large super hospital, of course,

385 00:21:02.070 --> 00:21:05.520 this is a hypothetical one but we can do that.

386 00:21:05.520 --> 00:21:10.197 And that is kind of like a national average thing, right?

387 00:21:10.197 --> 00:21:15.197 These are all reasonable ways to define a benchmark.

388 00:21:17.460 --> 00:21:19.470 And there is the last one.

389 00:21:19.470 --> 00:21:22.800 So the last one is kind of more like equity driven thing.

390 00:21:22.800 --> 00:21:25.620 So we can form a benchmark such that say,

391 00:21:25.620 --> 00:21:27.053 okay, say,

392 00:21:27.053 --> 00:21:29.100 from the regulator's perspective,

393 00:21:29.100 --> 00:21:33.780 we really wanna push hospitals to improve their performance

394 00:21:33.780 --> 00:21:35.760 for minority patients.

395 00:21:35.760 --> 00:21:40.760 So say, we can set the benchmark to be something like,

396 00:21:41.100 --> 00:21:43.230 okay, for within the minority groups,

397 00:21:43.230 --> 00:21:48.230 we can intentionally select patients with better outcomes.

398 00:21:48.426 --> 00:21:51.030 We can make the proportion to be very large

399 00:21:51.030 --> 00:21:54.450 so that in the benchmark group,

400 00:21:54.450 --> 00:21:59.130 we can have a very good performance for minority patients.

401 00:21:59.130 --> 00:22:02.880 And then black non-Hispanic patients.

402 00:22:02.880 --> 00:22:06.300 So this is kind of a equity driven thing.

403 00:22:06.300 --> 00:22:08.733 So as you can see, I give you like,

404 00:22:10.500 --> 00:22:12.450 at least the four examples.

405 00:22:12.450 --> 00:22:15.420 But these are scattered in the literature

406 00:22:15.420 --> 00:22:17.730 and there is no unifying framework

407 00:22:17.730 --> 00:22:20.220 to accommodate all of these cases.

408 00:22:20.220 --> 00:22:24.840 But we actually can develop a general framework

409 00:22:24.840 --> 00:22:26.396 to accommodate all.

410 00:22:26.396 --> 00:22:29.940 I will give you the details later.

411 00:22:29.940 --> 00:22:31.653 So, all right,

412 00:22:33.510 --> 00:22:36.390 so the framework

413 00:22:36.390 --> 00:22:39.960 that we proposed is what we termed,

414 00:22:39.960 --> 00:22:42.570 a versatile deep learning provider profiling.

415 00:22:42.570 --> 00:22:47.570 So we proposed a versatile or probabilistic framework

416 00:22:49.740 --> 00:22:51.900 based on the, so-called provider comparators,

417 00:22:51.900 --> 00:22:55.740 which is, you can name it as you know, provider comparator,

418 00:22:55.740 --> 00:22:58.050 hypothetical provider performance benchmark

419 00:22:58.050 --> 00:22:59.280 or population norm.

420 00:22:59.280 --> 00:23:02.880 These are all the same interchangeable terms.

421 00:23:02.880 --> 00:23:04.020 Okay?

422 00:23:04.020 --> 00:23:05.610 Here versatile means, okay,

423 00:23:05.610 --> 00:23:09.990 we can use the framework to do a lot of different things.

424 00:23:09.990 --> 00:23:13.590 So they are adaptable to different profiling objectives

425 00:23:13.590 --> 00:23:14.929 and contexts, okay?

426 00:23:14.929 --> 00:23:18.330 It's why we use the term versatile

427 00:23:18.330 --> 00:23:20.820 and here provider comparator,

428 00:23:20.820 --> 00:23:25.420 which is defined to be a hypothetical reference provider

429 00:23:27.513 --> 00:23:30.270 that is corresponding to your profiling objective.

430 00:23:30.270 --> 00:23:32.280 So if you have a certain objective,

431 00:23:32.280 --> 00:23:36.630 of course you can define your own hypothetical provider.

432 00:23:36.630 --> 00:23:39.150 And if you have a different objective,

433 00:23:39.150 --> 00:23:41.670 you can define another one, okay?

434 00:23:41.670 --> 00:23:44.520 And the deep learning thing comes

435 00:23:44.520 --> 00:23:48.660 into play because it is nice that,

436 00:23:48.660 --> 00:23:51.046 generally it relaxed the linearity assumption

437 00:23:51.046 --> 00:23:54.870 in most existing portfolio models

438 00:23:54.870 --> 00:23:57.930 that relies heavily on linear this assumption.

439 00:23:57.930 --> 00:23:58.763 Okay?

440 00:24:00.030 --> 00:24:05.030 Alright, so this is slide of the basic setup

441 00:24:07.410 --> 00:24:09.480 of this new approach.

442 00:24:09.480 --> 00:24:12.990 So let's say we have a ID random sample

443 00:24:12.990 --> 00:24:17.310 with Y as the outcome,

444 00:24:17.310 --> 00:24:21.960 and the Fi star is the provider identifier,

445 00:24:21.960 --> 00:24:26.043 and Zi is simply a vector of variants,

446 00:24:27.270 --> 00:24:31.593 and they are one from a population Y, F star, Z.

447 00:24:37.955 --> 00:24:39.720 And we have the following assumptions

448 00:24:39.720 --> 00:24:42.903 that these two assumptions, one and two,

449 00:24:46.351 --> 00:24:47.370 so F star.

450 00:24:47.370 --> 00:24:51.992 So basically this script F star is the support

451 00:24:51.992 --> 00:24:56.315 of this provider identifier, F star.

452 00:24:56.315 --> 00:24:57.796 Okay?

453 00:24:57.796 --> 00:25:02.796 So we require that this report for any value

454 00:25:04.770 --> 00:25:06.840 that this F star can pay,

455 00:25:06.840 --> 00:25:11.010 we assume that the probability of F star equal

456 00:25:11.010 --> 00:25:12.810 to F is positive,

457 00:25:12.810 --> 00:25:14.820 which means that in the dataset,

458 00:25:14.820 --> 00:25:19.020 you can at least observe one patient from that provider.

459 00:25:19.020 --> 00:25:19.950 Okay?

460 00:25:19.950 --> 00:25:23.640 Say if this is zero, then basically it means,

461 00:25:23.640 --> 00:25:27.307 okay, we do not observe any patient from that provider,

462 00:25:27.307 --> 00:25:29.103 which is useless, right?

463 00:25:30.960 --> 00:25:34.410 So the second assumption is simply,

464 00:25:34.410 --> 00:25:39.410 okay, so this script F star includes all possible providers,

465 00:25:41.310 --> 00:25:42.360 we wanna evaluate.

466 00:25:42.360 --> 00:25:45.180 So basically this F star has to fall

467 00:25:45.180 --> 00:25:48.630 into this set of values, okay?

468 00:25:48.630 --> 00:25:52.297 So that's why it's the probability as equal to one.

469 00:25:52.297 --> 00:25:53.130 Okay?

470 00:25:54.210 --> 00:25:58.350 So we have two important assumptions,

471 00:25:58.350 --> 00:26:00.210 regarding data generating mechanism.

472 00:26:00.210 --> 00:26:02.820 So the first one is basically the distribution

473 00:26:02.820 --> 00:26:04.800 of this F star.

474 00:26:04.800 --> 00:26:09.800 The provider identifier depends on covariate.

475 00:26:10.110 --> 00:26:14.070 And this is like, okay, so for a patient,

476 00:26:14.070 --> 00:26:17.130 say, I'm a patient, I wanna choose my provider,

477 00:26:17.130 --> 00:26:19.260 I wanna choose my hospital,

478 00:26:19.260 --> 00:26:21.150 my decision will largely based on,

479 00:26:21.150 --> 00:26:23.490 okay, what conditions I have,

480 00:26:23.490 --> 00:26:26.970 and what insurance I have, right?

481 00:26:26.970 --> 00:26:31.320 And say what is the possible feasible set

482 00:26:31.320 --> 00:26:33.570 of hospitals I can choose from?

483 00:26:33.570 --> 00:26:34.403 Okay?

484 00:26:34.403 --> 00:26:35.673 So these are all covariates

485 00:26:35.673 --> 00:26:37.350 that we can include in the model.

486 00:26:37.350 --> 00:26:41.220 So basically the F star is the distribution

487 00:26:41.220 --> 00:26:44.910 of a star depends on all those covariates

488 00:26:44.910 --> 00:26:47.730 which is reasonable assumption.

489 00:26:47.730 --> 00:26:48.780 The second one,

490 00:26:48.780 --> 00:26:50.910 the distribution of the outcome Y

491 00:26:50.910 --> 00:26:54.360 as a function of Z and F star,

492 00:26:54.360 --> 00:26:57.191 which means that, okay, the outcome,

493 00:26:57.191 --> 00:27:02.191 if I go to the hospital and say I have a certain disease

494 00:27:03.150 --> 00:27:08.150 and I got a treatment and whether I feel better

495 00:27:08.280 --> 00:27:09.990 or not really depends on, okay,

496 00:27:09.990 --> 00:27:11.910 of course, depends on my conditions,

497 00:27:11.910 --> 00:27:15.873 and also depends on which hospitals I went to, right?

498 00:27:17.100 --> 00:27:20.310 So the distribution is denoted

499 00:27:20.310 --> 00:27:24.870 as pi, y, given Z and F star.

500 00:27:24.870 --> 00:27:26.340 Okay?

501 00:27:26.340 --> 00:27:30.550 So basically these two assumptions gives us the,

502 00:27:30.550 --> 00:27:35.010 basically the basic setting for a patient who is looking

503 00:27:35.010 --> 00:27:39.783 for care to improve their conditions.

504 00:27:42.450 --> 00:27:47.450 So the main idea in this new framework is reclassification.

505 00:27:47.910 --> 00:27:48.743 So basically,

506 00:27:48.743 --> 00:27:53.743 we wanna construct a hypothetical provider comparator

507 00:27:53.970 --> 00:27:55.950 as a performance benchmark

508 00:27:55.950 --> 00:28:00.950 that is corresponding to our specific profiling objective.

509 00:28:01.363 --> 00:28:02.196 Okay?

510 00:28:02.196 --> 00:28:05.880 So reclassification here means that we wanna,

511 00:28:05.880 --> 00:28:10.680 we reclassify subjects from existing providers

512 00:28:10.680 --> 00:28:12.900 into a hypothetical one

513 00:28:12.900 --> 00:28:15.270 following a certain probability distribution.

514 00:28:15.270 --> 00:28:16.110 Okay?

515 00:28:16.110 --> 00:28:18.930 To do this, we introduced a random indicator,

516 00:28:18.930 --> 00:28:20.790 it's just a 0, 1.

517 00:28:20.790 --> 00:28:23.650 Which we termed reclassifier.

518 00:28:23.650 --> 00:28:26.400 This reclassifier is equal to 0.

519 00:28:26.400 --> 00:28:27.840 Here it is kind of different.

520 00:28:27.840 --> 00:28:30.994 So reclassifier is equal to zero.

521 00:28:30.994 --> 00:28:33.446 When the subject is reclassified

522 00:28:33.446 --> 00:28:35.053 into the hypothetical provider,

523 00:28:35.053 --> 00:28:39.326 if it is equal to one, then the subject is not reclassified.

524 00:28:39.326 --> 00:28:43.826 So the patient stays in their original provider, okay?

525 00:28:46.590 --> 00:28:50.610 And with this reclassified redefined, F,

526 00:28:50.610 --> 00:28:52.953 so F is different from F star.

527 00:28:53.850 --> 00:28:57.180 So F is defined as the product of R,

528 00:28:57.180 --> 00:28:59.250 basically R times F star.

529 00:28:59.250 --> 00:29:04.250 And we basically add a singleton to this F script F star.

530 00:29:05.910 --> 00:29:09.387 So now we can see, okay,

531 00:29:09.387 --> 00:29:13.230 so whatever providers we have originally,

532 00:29:13.230 --> 00:29:16.320 now we add a single hypothetical provider

533 00:29:16.320 --> 00:29:21.300 and we provide the provider indicator,

534 00:29:21.300 --> 00:29:23.310 we fix that as zero.

535 00:29:23.310 --> 00:29:25.770 So zero is the hypothetical one.

536 00:29:25.770 --> 00:29:29.430 So now this F can take values,

537 00:29:29.430 --> 00:29:31.350 importantly, it can take whatever values

538 00:29:31.350 --> 00:29:33.870 from the original script F

539 00:29:33.870 --> 00:29:37.320 but now it can also take values

540 00:29:37.320 --> 00:29:40.020 to take the value zero, right?

541 00:29:40.020 --> 00:29:42.840 So basically this R is used

542 00:29:42.840 --> 00:29:46.290 to manipulate a subject's provider membership.

543 00:29:46.290 --> 00:29:51.290 So, a subject from a provider F star equal to F.

544 00:29:53.850 --> 00:29:55.170 So here in this case,

545 00:29:55.170 --> 00:29:58.590 because it's F star, it cannot be equal to zero, right?

546 00:29:58.590 --> 00:30:00.660 So we wanna reclassify patients

547 00:30:00.660 --> 00:30:03.750 from a certain existing real provider

548 00:30:03.750 --> 00:30:05.793 to that hypothetical provider.

549 00:30:07.172 --> 00:30:09.960 You know, this F is equal to zero.

550 00:30:09.960 --> 00:30:13.761 So this is a new provider membership for that patient, okay?

551 00:30:13.761 --> 00:30:15.810 But if R is equal to zero,

552 00:30:15.810 --> 00:30:19.890 then the patient stays in that original hospital.

553 00:30:19.890 --> 00:30:20.723 Okay?

554 00:30:22.230 --> 00:30:23.280 Alright.

555 00:30:23.280 --> 00:30:24.900 We have additional two assumptions

556 00:30:24.900 --> 00:30:28.770 regarding this reclassification thing.

557 00:30:28.770 --> 00:30:33.770 So the first one is for any provider, real provider,

558 00:30:35.070 --> 00:30:38.400 we have this probability, being less than one.

559 00:30:38.400 --> 00:30:40.140 This means that, okay,

560 00:30:40.140 --> 00:30:44.340 so given a set of covariates and given

561 00:30:44.340 --> 00:30:48.993 that the patient is in a certain provider,

562 00:30:48.993 --> 00:30:52.740 then the patient being reclassified

563 00:30:52.740 --> 00:30:55.560 into the new hypothetical provider,

564 00:30:55.560 --> 00:30:57.900 the probability is less than one,

565 00:30:57.900 --> 00:31:02.900 which means that we should keep at least a few patients

566 00:31:03.240 --> 00:31:05.040 in their original provider

567 00:31:05.040 --> 00:31:09.750 so that we can still evaluate the outcome distributions

568 00:31:09.750 --> 00:31:11.673 of the original provider, okay?

569 00:31:13.298 --> 00:31:15.030 And this actually,

570 00:31:15.030 --> 00:31:18.543 if you do some, a simple algebra,

571 00:31:20.292 --> 00:31:23.490 we can show that basically this implies that,

572 00:31:23.490 --> 00:31:25.560 I mean this, we can basically drop this condition

573 00:31:25.560 --> 00:31:27.840 because if you do the sum

574 00:31:27.840 --> 00:31:31.260 of the conditional probability thing,

575 00:31:31.260 --> 00:31:33.000 you can basically drop this condition

576 00:31:33.000 --> 00:31:34.830 and this actually holds.

577 00:31:34.830 --> 00:31:38.130 So it's like, okay, no matter which hospital,

578 00:31:38.130 --> 00:31:42.030 no matter which provider the patient is in currently,

579 00:31:42.030 --> 00:31:43.650 the probability that the patient

580 00:31:43.650 --> 00:31:46.020 will be reclassified is less than one.

581 00:31:46.020 --> 00:31:49.920 So not all patients will be reclassified, right?

582 00:31:49.920 --> 00:31:52.080 And this is the second condition.

583 00:31:52.080 --> 00:31:55.733 So combining these two, basically, okay,

584 00:31:58.170 --> 00:32:01.740 so basically not all patients can be reclassified

585 00:32:01.740 --> 00:32:05.603 or also all patients cannot be living

586 00:32:07.920 --> 00:32:10.020 in their original providers.

587 00:32:10.020 --> 00:32:14.580 Basically we require that, okay, each patient can,

588 00:32:14.580 --> 00:32:16.830 so we should have

589 00:32:16.830 --> 00:32:19.920 at least a few patients who are remaining

590 00:32:19.920 --> 00:32:22.020 in their original hospitals so that we can evaluate

591 00:32:22.020 --> 00:32:24.600 their original outcome distributions.

592 00:32:24.600 --> 00:32:28.290 And also we need a, of course characterize the distribution,

593 00:32:28.290 --> 00:32:31.410 that hypothetical reference provider.

594 00:32:31.410 --> 00:32:32.243 Okay?

595 00:32:33.360 --> 00:32:34.380 Alright.

596 00:32:34.380 --> 00:32:37.793 Then the last assumption is,

597 00:32:37.793 --> 00:32:40.230 this is kind of an interesting setting.

598 00:32:40.230 --> 00:32:43.830 So rather than observing the original data,

599 00:32:43.830 --> 00:32:48.213 Y, F star, Z, we can only observe this set.

600 00:32:51.540 --> 00:32:56.513 So it's R, Y, F, Z, this tuple.

601 00:32:58.810 --> 00:33:02.010 So the big difference between these two is,

602 00:33:02.010 --> 00:33:05.370 for this one, we know exactly for every patient,

603 00:33:05.370 --> 00:33:08.040 we know exactly where they're from,

604 00:33:08.040 --> 00:33:11.040 which provider they are in.

605 00:33:11.040 --> 00:33:15.170 But for the this one, say if R is equal to 0,

606 00:33:16.582 --> 00:33:18.270 F is automatically 0

607 00:33:18.270 --> 00:33:20.913 because F is defined as R times F star.

608 00:33:21.780 --> 00:33:23.400 So for those patients,

609 00:33:23.400 --> 00:33:26.673 we actually don't know where they come from, right?

610 00:33:27.510 --> 00:33:30.510 But here we assume

611 00:33:30.510 --> 00:33:34.680 that we can only observe post-reclassification data.

612 00:33:34.680 --> 00:33:37.200 And this actually is nice,

613 00:33:37.200 --> 00:33:42.200 I mean this is not always necessary in the practice,

614 00:33:42.300 --> 00:33:45.480 but this assumption actually helps,

615 00:33:45.480 --> 00:33:49.500 facilitates the implementation

616 00:33:49.500 --> 00:33:52.830 of some certain privacy preserving protocols

617 00:33:52.830 --> 00:33:54.000 and data security protocols.

618 00:33:54.000 --> 00:33:56.193 If say, okay, we don't want the,

619 00:33:57.300 --> 00:34:00.540 because of certain powerful influential providers

620 00:34:00.540 --> 00:34:05.100 can actually have a strong influence

621 00:34:05.100 --> 00:34:06.660 in policy making.

622 00:34:06.660 --> 00:34:10.560 So, because this is capped like confidential,

623 00:34:10.560 --> 00:34:14.010 so they actually don't know how we design,

624 00:34:14.010 --> 00:34:17.610 how we choose the re-classification scheme.

625 00:34:17.610 --> 00:34:22.610 So it can help reduce some unwarranted inference

626 00:34:24.930 --> 00:34:28.800 from those very powerful stakeholders.

627 00:34:28.800 --> 00:34:31.530 So this is a nice setting,

628 00:34:31.530 --> 00:34:34.563 but it doesn't have to be like this in reality.

629 00:34:36.120 --> 00:34:40.830 Alright, so now we have four assumptions,

630 00:34:40.830 --> 00:34:41.760 important assumptions

631 00:34:41.760 --> 00:34:43.920 to regarding the data generating mechanism

632 00:34:43.920 --> 00:34:47.670 and to regarding the reclassification scheme.

633 00:34:47.670 --> 00:34:52.670 So, the ultimate goals of profiling is

634 00:34:54.180 --> 00:34:56.670 to first to evaluate all providers,

635 00:34:56.670 --> 00:34:58.953 and then we wanna identify goals,

636 00:35:00.270 --> 00:35:02.760 especially with very bad performance

637 00:35:02.760 --> 00:35:06.750 and we can take additional actions

638 00:35:06.750 --> 00:35:08.943 and so we can, you know,

639 00:35:09.870 --> 00:35:11.880 improve their performance in certain way.

640 00:35:11.880 --> 00:35:12.713 Okay?

641 00:35:12.713 --> 00:35:13.920 But yeah,

642 00:35:13.920 --> 00:35:17.583 so this quantitatively or mathematically,

643 00:35:20.314 --> 00:35:22.140 we have the two overarching goals.

644 00:35:22.140 --> 00:35:24.450 The first one is to harness,

645 00:35:24.450 --> 00:35:28.830 to use the post reclassification data,

646 00:35:28.830 --> 00:35:32.890 to contrast the distribution of each existing

647 00:35:35.360 --> 00:35:36.960 or real provider.

648 00:35:36.960 --> 00:35:41.493 F star was the newly defined reference group.

649 00:35:42.450 --> 00:35:43.560 So we wanna compare,

650 00:35:43.560 --> 00:35:46.950 basically, compare the distribution of these two groups.

651 00:35:46.950 --> 00:35:48.300 I mean each of them

652 00:35:48.300 --> 00:35:50.850 because we have so many real providers,

653 00:35:50.850 --> 00:35:53.790 and we only have a single hypothetical provider, okay?

654 00:35:53.790 --> 00:35:56.940 We wanna compare them, we wanna do contrasts.

655 00:35:56.940 --> 00:35:58.830 And of course the second goal is

656 00:35:58.830 --> 00:36:03.830 to identify those providers with very bad performance.

657 00:36:06.420 --> 00:36:07.253 All right,

658 00:36:08.400 --> 00:36:09.933 so, this actually,

659 00:36:14.040 --> 00:36:16.890 because we introduced this hypothetical provider,

660 00:36:16.890 --> 00:36:18.600 this is really nice actually.

661 00:36:18.600 --> 00:36:23.600 But there is a difficult issue here

662 00:36:23.670 --> 00:36:27.967 because we introduced this hypothetical provider,

663 00:36:29.850 --> 00:36:31.950 we actually have to account for

664 00:36:31.950 --> 00:36:35.070 or address reclassification dues to bias.

665 00:36:35.070 --> 00:36:39.240 So the details are in this proposition.

666 00:36:39.240 --> 00:36:42.510 So let's assume that those four assumptions hold

667 00:36:42.510 --> 00:36:47.327 and the distribution of the outcome given Z and this F,

668 00:36:50.070 --> 00:36:52.713 F is the newly defined provider indicator.

669 00:36:53.850 --> 00:36:58.200 We can actually write the outcome distribution,

670 00:36:58.200 --> 00:36:59.490 like in two cases.

671 00:36:59.490 --> 00:37:01.468 So when F is equal to 0,

672 00:37:01.468 --> 00:37:04.748 this is corresponding to the reference,

673 00:37:04.748 --> 00:37:06.502 the hypothetical provider.

674 00:37:06.502 --> 00:37:09.169 So this is actually the average,

675 00:37:13.609 --> 00:37:17.859 you can consider as the distribution of the outcome

676 00:37:20.760 --> 00:37:22.080 basically for all patient.

677 00:37:22.080 --> 00:37:24.903 If you group all patients together into a single group,

678 00:37:24.903 --> 00:37:28.440 this is basically the distribution of that group.

679 00:37:28.440 --> 00:37:29.273 Okay?

680 00:37:29.273 --> 00:37:31.470 But we have this term here,

681 00:37:31.470 --> 00:37:34.410 and this is not necessarily equal to 1,

682 00:37:34.410 --> 00:37:37.920 F is equal to 1 then it's very simple,

683 00:37:37.920 --> 00:37:42.920 but it could be unequal to 1.

684 00:37:43.290 --> 00:37:47.970 And also in the second case when F is not equal to 0,

685 00:37:47.970 --> 00:37:51.663 which means that okay, for those existing providers,

686 00:37:52.740 --> 00:37:55.830 their distribution also changes because you basically,

687 00:37:55.830 --> 00:37:59.490 you move a few patients to the new provider.

688 00:37:59.490 --> 00:38:02.490 So the original distribution changes, right?

689 00:38:02.490 --> 00:38:05.880 And because we cannot observe this by assumption.

690 00:38:05.880 --> 00:38:09.900 So this is basically the observed outcome distribution

691 00:38:09.900 --> 00:38:11.490 for existing providers.

692 00:38:11.490 --> 00:38:13.620 But according, as you can see here,

693 00:38:13.620 --> 00:38:14.640 it's a bias distribution.

694 00:38:14.640 --> 00:38:16.770 It's no longer the original one, right?

695 00:38:16.770 --> 00:38:18.060 Because this ratio, again,

696 00:38:18.060 --> 00:38:20.793 it is not necessarily equal to 1, okay?

697 00:38:22.320 --> 00:38:23.153 Right?

698 00:38:23.153 --> 00:38:25.629 So as I said,

699 00:38:25.629 --> 00:38:28.920 you can consider this as the average distribution,

700 00:38:28.920 --> 00:38:30.960 basically as the outcome distribution

701 00:38:30.960 --> 00:38:33.150 of the whole patient population, okay?

702 00:38:33.150 --> 00:38:38.150 So of course you can write it as a sum of the,

703 00:38:38.700 --> 00:38:40.863 you know, weighted probabilities.

704 00:38:42.390 --> 00:38:46.080 So the weight being the probability provider membership,

705 00:38:46.080 --> 00:38:50.280 and this is basically, okay, within this certain provider,

706 00:38:50.280 --> 00:38:52.830 what does the outcome distribution look like?

707 00:38:52.830 --> 00:38:56.043 Okay. All right.

708 00:38:57.600 --> 00:38:59.013 So a few things.

709 00:39:02.670 --> 00:39:06.030 This proposition basically outlines a,

710 00:39:06.030 --> 00:39:08.130 what we call design based approach

711 00:39:08.130 --> 00:39:11.910 to provider profiling, basically, okay.

712 00:39:11.910 --> 00:39:12.903 So,

713 00:39:14.430 --> 00:39:17.130 I actually, I mentioned this early,

714 00:39:17.130 --> 00:39:19.830 in profiling there are a few different parties.

715 00:39:19.830 --> 00:39:22.740 The first one is regulars

716 00:39:22.740 --> 00:39:25.080 who initiated the profiling process

717 00:39:25.080 --> 00:39:25.950 because they are interested

718 00:39:25.950 --> 00:39:28.200 in the performance of these providers.

719 00:39:28.200 --> 00:39:30.570 And also we have profilers,

720 00:39:30.570 --> 00:39:33.330 which basically evaluates the performance,

721 00:39:33.330 --> 00:39:36.390 but they don't have to be the same as regulators.

722 00:39:36.390 --> 00:39:37.740 And also we have of course,

723 00:39:37.740 --> 00:39:41.550 providers who are the subject of evaluation

724 00:39:41.550 --> 00:39:43.950 and we also have patient who need the information

725 00:39:43.950 --> 00:39:45.840 to make their decision, okay?

726 00:39:45.840 --> 00:39:47.220 So the design-based approach

727 00:39:47.220 --> 00:39:51.150 basically tell us that, okay, so, for regulators,

728 00:39:51.150 --> 00:39:53.280 they can basically lead the development

729 00:39:53.280 --> 00:39:56.790 of a reclassification scene because in this framework,

730 00:39:56.790 --> 00:39:59.460 we never say what the distribution,

731 00:39:59.460 --> 00:40:02.100 say, what this looks like, where, right?

732 00:40:02.100 --> 00:40:05.370 So this is a very general specification

733 00:40:05.370 --> 00:40:08.430 and we only made that four assumptions,

734 00:40:08.430 --> 00:40:12.150 but we don't have any distributional assumption here.

735 00:40:12.150 --> 00:40:15.270 So we can make it very general.

736 00:40:15.270 --> 00:40:18.870 And so in this framework,

737 00:40:18.870 --> 00:40:23.580 regulators will get more involved in this process.

738 00:40:23.580 --> 00:40:25.350 So that's why they can

739 00:40:25.350 --> 00:40:30.000 basically design the reclassification scheme

740 00:40:30.000 --> 00:40:33.303 based on their specific objectives, okay?

741 00:40:34.650 --> 00:40:35.483 Alright.

742 00:40:35.483 --> 00:40:40.483 So, and given a specific reclassification scheme,

743 00:40:40.500 --> 00:40:44.550 of course they can design their own reference group,

744 00:40:44.550 --> 00:40:47.160 their hypothetical providers

745 00:40:47.160 --> 00:40:51.900 and having defined this hypothetical provider,

746 00:40:51.900 --> 00:40:56.350 profilers of course can use post the reclassification data

747 00:41:00.090 --> 00:41:01.230 and also the dependence.

748 00:41:01.230 --> 00:41:03.240 Because here, as you can see here,

749 00:41:03.240 --> 00:41:04.890 this R actually depends on Y,

750 00:41:04.890 --> 00:41:06.390 depends on the outcome covariate

751 00:41:06.390 --> 00:41:09.870 and the provider identification.

752 00:41:09.870 --> 00:41:12.337 So using this information

753 00:41:16.350 --> 00:41:18.633 and also the post reclassification data,

754 00:41:19.620 --> 00:41:23.250 profilers that can actually do the profiling

755 00:41:23.250 --> 00:41:26.070 and we can use the framework

756 00:41:26.070 --> 00:41:29.310 to estimate the probabilities reclassification,

757 00:41:29.310 --> 00:41:32.550 which is also the propensity scores actually.

758 00:41:32.550 --> 00:41:34.900 So the next step would be

759 00:41:38.790 --> 00:41:40.830 to use the estimated propensity scores

760 00:41:40.830 --> 00:41:45.270 to correct for reclassification induced bias.

761 00:41:45.270 --> 00:41:50.270 And then we can basically construct the distribution

762 00:41:51.390 --> 00:41:55.470 of the hypothetical provider with the distribution

763 00:41:55.470 --> 00:41:58.173 of the existing provider, okay?

764 00:41:59.970 --> 00:42:01.020 Alright.

765 00:42:01.020 --> 00:42:05.490 So as sketched in the previous slide,

766 00:42:05.490 --> 00:42:07.410 there are a few important things

767 00:42:07.410 --> 00:42:11.070 or advantages of the design-based approach.

768 00:42:11.070 --> 00:42:12.570 So this approach actually,

769 00:42:12.570 --> 00:42:14.160 in this framework,

770 00:42:14.160 --> 00:42:18.633 providers can be more involved in this framework.

771 00:42:21.060 --> 00:42:21.893 And,

772 00:42:23.730 --> 00:42:28.020 so we can use the profiling result,

773 00:42:28.020 --> 00:42:30.310 from this new approach can be more relevant

774 00:42:31.320 --> 00:42:32.730 to what people are interested

775 00:42:32.730 --> 00:42:37.083 in the care decision making process, okay?

776 00:42:37.970 --> 00:42:42.721 So, I think I'm a bit over time,

777 00:42:42.721 --> 00:42:47.721 but I wanna quickly skim through a few examples.

778 00:42:47.760 --> 00:42:51.030 But these examples are basically,

779 00:42:51.030 --> 00:42:53.610 we need a few assumptions like

780 00:42:53.610 --> 00:42:57.843 whether the reclassifier is depending on the outcome,

781 00:42:59.790 --> 00:43:02.190 so in this example, it’s very simple.

782 00:43:02.190 --> 00:43:07.190 Basically the reclassifier is independent of everything.

783 00:43:07.500 --> 00:43:08.640 So,

784 00:43:08.640 --> 00:43:11.520 actually this reduces to the most simple case.

785 00:43:11.520 --> 00:43:15.330 So nothing changes actually after reclassification,

786 00:43:15.330 --> 00:43:20.023 but this is an example about the setting.

787 00:43:21.360 --> 00:43:25.833 And we also have like a few non-dependent settings.

788 00:43:27.300 --> 00:43:32.220 This R can depend on F star and given F star,

789 00:43:32.220 --> 00:43:34.830 it can be independent with Y.

790 00:43:34.830 --> 00:43:36.600 And we also have some examples,

791 00:43:36.600 --> 00:43:39.420 this is called equal rate representation.

792 00:43:39.420 --> 00:43:41.850 We also have singular representation,

793 00:43:41.850 --> 00:43:42.900 basically the setting

794 00:43:42.900 --> 00:43:45.130 where we only choose a single provider

795 00:43:46.956 --> 00:43:48.540 and we also have the case

796 00:43:48.540 --> 00:43:53.540 where R actually depends on Y, the outcome.

797 00:43:54.270 --> 00:43:56.430 So we can basically choose the outcome,

798 00:43:56.430 --> 00:44:01.430 sorry, we can choose patients based on the outcome.

799 00:44:02.280 --> 00:44:04.050 And I also give an example,

800 00:44:04.050 --> 00:44:05.880 this is actually an interesting example,

801 00:44:05.880 --> 00:44:09.600 but seems like we don’t have enough time today.

802 00:44:09.600 --> 00:44:13.470 So this is the most general case where R is allowed

803 00:44:13.470 --> 00:44:17.880 to depend on F and also Y.

804 00:44:17.880 --> 00:44:19.740 So we don’t have independence anymore,

805 00:44:19.740 --> 00:44:21.810 but unfortunately this case,

806 00:44:21.810 --> 00:44:25.830 we have the unidentifiability issue.

807 00:44:25.830 --> 00:44:27.210 So this case won't work

808 00:44:27.210 --> 00:44:29.643 under the post-reclassification data assumption.

809 00:44:30.840 --> 00:44:34.803 So we actually developed a framework,

810 00:44:37.560 --> 00:44:40.530 we looked at the deep learning methods

811 00:44:40.530 --> 00:44:44.220 and the singular representation case.

812 00:44:44.220 --> 00:44:46.710 And this is a relatively simple framework.

813 00:44:46.710 --> 00:44:50.430 We only consider exponential distribution.

814 00:44:50.430 --> 00:44:51.263 I mean the outcome

815 00:44:51.263 --> 00:44:54.280 involves the exponential family distribution

816 00:44:56.100 --> 00:44:59.193 and we construct a neural network model.

817 00:45:00.810 --> 00:45:02.277 So we have the input layer

818 00:45:02.277 --> 00:45:05.490 and the fully connected hidden layers and the outcome layer,

819 00:45:05.490 --> 00:45:10.490 and we use stratify sampling based optimization algorithm.

820 00:45:11.070 --> 00:45:14.490 Here, I will skip the detail.

821 00:45:14.490 --> 00:45:19.380 And we developed a exact test based outcome distribution,

822 00:45:19.380 --> 00:45:24.380 exact test based approach to identify outlined performers.

823 00:45:25.410 --> 00:45:26.520 Okay?

824 00:45:26.520 --> 00:45:29.010 And this is basically the motivation

825 00:45:29.010 --> 00:45:32.490 why we need deep learning here, because simply speaking,

826 00:45:32.490 --> 00:45:35.700 the covid effect is not constant over calendar time

827 00:45:35.700 --> 00:45:39.060 and we have to easily account for that

828 00:45:39.060 --> 00:45:40.380 while doing profiling,

829 00:45:40.380 --> 00:45:43.023 but the effect itself is not of interest.

830 00:45:47.677 --> 00:45:50.793 Basically a visualization of the profile results.

831 00:45:52.920 --> 00:45:55.593 So here we construct the,

832 00:45:57.000 --> 00:46:01.350 we construct what we call the funnel plot here.

833 00:46:01.350 --> 00:46:05.520 So the benchmark, the reference, the indicator,

834 00:46:05.520 --> 00:46:09.180 we use is again Oi divided by Ei

835 00:46:09.180 --> 00:46:14.180 and Ei and defined where this one is the median.

836 00:46:14.280 --> 00:46:18.330 And this is actually the neural network part.

837 00:46:18.330 --> 00:46:21.750 And we have the funnel plots.

838 00:46:21.750 --> 00:46:25.470 So those dots represent providers, okay?

839 00:46:25.470 --> 00:46:27.483 So because this, I mean,

840 00:46:28.641 --> 00:46:30.560 the higher, the worse the performance,

841 00:46:30.560 --> 00:46:32.280 the lower, the better the performance.

842 00:46:32.280 --> 00:46:36.933 So these blue dots here are actually better performers.

843 00:46:37.920 --> 00:46:42.202 So as you can see, if you add these two supporters up,

844 00:46:42.202 --> 00:46:44.423 this is like over 20%,

845 00:46:47.010 --> 00:46:48.420 what does not make practical sense

846 00:46:48.420 --> 00:46:52.050 because in practice you cannot identify outliers

847 00:46:52.050 --> 00:46:56.760 with over 20%, you know, this is too much.

848 00:46:56.760 --> 00:46:58.890 So we have to somehow account

849 00:46:58.890 --> 00:47:02.940 for provider level unmeasured confounding.

850 00:47:02.940 --> 00:47:07.940 And I didn't include the technical details here.

851 00:47:08.070 --> 00:47:10.500 But after the adjustment,

852 00:47:10.500 --> 00:47:13.890 as you can see the proportion of a better

853 00:47:13.890 --> 00:47:17.790 and the worse performers are much lower than before.

854 00:47:17.790 --> 00:47:22.790 And I think I only have one more slide.

855 00:47:23.220 --> 00:47:25.110 So some takeaways.

856 00:47:25.110 --> 00:47:27.000 So profiling is very important

857 00:47:27.000 --> 00:47:30.390 as a major societal undertaking in the United States.

858 00:47:30.390 --> 00:47:33.150 And we have so many applications,

859 00:47:33.150 --> 00:47:37.860 important implications and important consequences as well.

860 00:47:37.860 --> 00:47:41.490 And the new framework actually

861 00:47:41.490 --> 00:47:44.730 increased the regulators engagement in this process.

862 00:47:44.730 --> 00:47:46.560 And it's called versatile

863 00:47:46.560 --> 00:47:49.230 because we can handle different profiling objectives

864 00:47:49.230 --> 00:47:50.310 and it is compatible

865 00:47:50.310 --> 00:47:52.560 with many different model specifications,

866 00:47:52.560 --> 00:47:54.900 machine learning models, data science models.

867 00:47:54.900 --> 00:47:57.810 And here we use deep learning

868 00:47:57.810 --> 00:48:01.090 because it relaxes the linearity assumption

869 00:48:01.090 --> 00:48:05.310 and it is often a good idea to account

870 00:48:05.310 --> 00:48:08.340 for provider level measure confounding

871 00:48:08.340 --> 00:48:10.740 when we do this profiling stuff.

872 00:48:10.740 --> 00:48:14.700 And that's all for today.

873 00:48:14.700 --> 00:48:15.873 Thank you so much for.

874 00:48:20.020 --> 00:48:22.020 I know we only have like two-

875 00:48:22.020 --> 00:48:23.340 <v Host>Yeah, We have two minutes.</v>

876 00:48:23.340 --> 00:48:26.310 Thank you very much Dr. Wu for your presentation.

877 00:48:26.310 --> 00:48:27.843 Any questions in the audience?

878 00:48:35.880 --> 00:48:36.960 Anyone online?

879 00:48:36.960 --> 00:48:38.460 Just giving everyone a chance.

880 00:48:39.960 --> 00:48:41.340 No, I'll ask a question.

881 00:48:41.340 --> 00:48:45.360 So I think it's really cool to be able

882 00:48:45.360 --> 00:48:47.910 to identify providers who are doing really well

883 00:48:47.910 --> 00:48:49.470 or doing bad.

884 00:48:49.470 --> 00:48:50.610 What do you do with that?

885 00:48:50.610 --> 00:48:51.900 Now that you have that result?

886 00:48:51.900 --> 00:48:53.970 Like do you tell the profiler

887 00:48:53.970 --> 00:48:55.863 or the patient get to give it to say,

888 00:48:55.863 --> 00:48:58.350 "Oh, I don't wanna go to them, they're bad."

889 00:48:58.350 --> 00:48:59.430 <v Dr. Wu>Yeah, that's a good question.</v>

890 00:48:59.430 --> 00:49:01.740 So actually CMS,

891 00:49:01.740 --> 00:49:06.740 they have many programs say, one is for dialysis patients,

892 00:49:06.930 --> 00:49:09.510 they have dialysis facility compare,

893 00:49:09.510 --> 00:49:11.640 which is an online program.

894 00:49:11.640 --> 00:49:15.120 So patient can have access to different types

895 00:49:15.120 --> 00:49:20.100 of information like whether diet facility is good or bad

896 00:49:20.100 --> 00:49:23.970 and many other different fields

897 00:49:23.970 --> 00:49:25.740 of information they have online.

898 00:49:25.740 --> 00:49:30.270 So they can choose their favorite providers.

899 00:49:30.270 --> 00:49:32.340 Yeah, that's possible.

900 00:49:32.340 --> 00:49:34.863 And it's something that is going on, yeah.

901 00:49:35.760 --> 00:49:37.013 <v Host>Oh, I think we have questions.</v>

902 00:49:37.013 --> 00:49:38.640 <v ->Yep.</v> <v ->Just very briefly,</v>

903 00:49:38.640 --> 00:49:40.490 because I know we're out of time but.

904 00:49:42.810 --> 00:49:44.550 To what extent do you feel that,

905 00:49:44.550 --> 00:49:47.460 if this is true, I guess, and doesn't matter,

906 00:49:47.460 --> 00:49:50.433 the patients don't necessarily have meetings.

907 00:49:51.420 --> 00:49:55.110 So for example, like I grew up in a rural county,

908 00:49:55.110 --> 00:49:57.090 we had one hospital, you were going to a hospital,

909 00:49:57.090 --> 00:49:58.230 you were going there.

910 00:49:58.230 --> 00:49:59.640 Even in New Haven,

911 00:49:59.640 --> 00:50:02.160 there are two campuses of Yale New Haven Hospital,

912 00:50:02.160 --> 00:50:06.030 but there's only one hospital in metro area.

913 00:50:06.030 --> 00:50:11.030 So, I mean, choice is kind of not a real thing.

914 00:50:11.160 --> 00:50:12.570 How does that affect?

915 00:50:12.570 --> 00:50:16.950 <v Dr. Wu>Right, that's a very good point, so-</v>

916 00:50:16.950 --> 00:50:17.790 <v Questioner>We are actually in city,</v>

917 00:50:17.790 --> 00:50:19.140 I understand there's more than one.

918 00:50:19.140 --> 00:50:21.130 (Host laughs) Right, there are so many.

919 00:50:21.130 --> 00:50:23.130 <v Dr. Wu>Yeah, but that's a very good point</v>

920 00:50:23.130 --> 00:50:27.030 because we are actually considering another framework

921 00:50:27.030 --> 00:50:29.670 which is also clustering framework,

922 00:50:29.670 --> 00:50:32.010 which basically gives you

923 00:50:32.010 --> 00:50:34.080 under certain conditions you can choose,

924 00:50:34.080 --> 00:50:36.000 there's a feasible set of providers

925 00:50:36.000 --> 00:50:37.290 that you can choose from,

926 00:50:37.290 --> 00:50:39.420 of course, under certain strengths,

927 00:50:39.420 --> 00:50:44.420 say your insurance, your location, many other conditions.

928 00:50:45.240 --> 00:50:47.793 But I mean, in this framework,

929 00:50:49.320 --> 00:50:51.130 maybe we can address that issue

930 00:50:52.590 --> 00:50:56.760 in the set of areas that we included here.

931 00:50:56.760 --> 00:51:01.760 But yeah, I mean, you know, very important issue.

932 00:51:05.602 --> 00:51:06.435 <v Host>Unfortunately, that's time.</v>

933 00:51:06.435 --> 00:51:08.768 So let's thank Dr. Wu again.

934 00:51:11.897 --> 00:51:15.027 If you haven't signed in, please sign in before you speak.

935 00:51:15.027 --> 00:51:16.773 You are registered.

936 00:51:16.773 --> 00:51:18.834 Oh no, it's good, I don't know.

937 00:51:18.834 --> 00:51:22.167 (indistinct chattering)