WEBVTT

1 00:00:02.010 --> 00:00:03.840 <v Maria>My name is Maria Ciarleglio</v>

2 00:00:03.840 --> 00:00:05.460 and I'm a faculty member

3 00:00:05.460 --> 00:00:07.740 in the Department of Biostatistics

4 00:00:07.740 --> 00:00:10.023 at the Yale School of Public Health.

5 00:00:10.890 --> 00:00:14.250 In this video series I will introduce the clinical research

6 00:00:14.250 --> 00:00:18.873 process to prepare you to collaborate with a statistician.

7 00:00:20.490 --> 00:00:23.550 In this first video we'll discuss what is often

8 00:00:23.550 --> 00:00:26.310 the first step of the research process,

9 00:00:26.310 --> 00:00:28.593 formulating a research question.

10 00:00:30.720 --> 00:00:32.880 The first step in the research process

11 00:00:32.880 --> 00:00:35.700 is to convert the need for information

12 00:00:35.700 --> 00:00:40.350 into an answerable question or hypothesis.

13 00:00:40.350 --> 00:00:45.350 A well formulated research question is specific and precise.

14 00:00:45.360 --> 00:00:48.540 The research question guides the study design

15 00:00:48.540 --> 00:00:52.200 and other design-related study characteristics,

16 00:00:52.200 --> 00:00:55.920 the data that are collected, the data analysis

17 00:00:55.920 --> 00:00:58.830 and ultimately determines what you can conclude

18 00:00:58.830 --> 00:01:00.303 at the end of the study.

19 00:01:02.250 --> 00:01:05.310 The PICO criteria can be used to guide you

20 00:01:05.310 --> 00:01:08.910 in framing a comparative research question.

21 00:01:08.910 --> 00:01:12.450 The PICO framework begins by specifying the population

22 00:01:12.450 --> 00:01:16.260 of interest, then the intervention being studied,

23 00:01:16.260 --> 00:01:19.170 the control or comparator group,

24 00:01:19.170 --> 00:01:21.183 and the outcomes of interest.

25 00:01:23.340 --> 00:01:26.280 Begin by specifying the population of interest.

26 00:01:26.280 --> 00:01:27.360 For example,

27 00:01:27.360 --> 00:01:31.290 patients with non-alcoholic fatty liver disease.

28 00:01:31.290 --> 00:01:34.050 The target population is the group of patients

29 00:01:34.050 --> 00:01:38.460 to which you would like to generalize your study findings.

30 00:01:38.460 --> 00:01:40.350 The study population is the group

31 00:01:40.350 --> 00:01:43.053 of patients to which you have access.

32 00:01:44.340 --> 00:01:47.370 The study population may be a subset

33 00:01:47.370 --> 00:01:49.293 of the target population.

34 00:01:50.190 --> 00:01:52.980 For example, your goal may be to generalize

35 00:01:52.980 --> 00:01:54.990 to all adult Americans

36 00:01:54.990 --> 00:01:58.290 with non-alcoholic fatty liver disease.

37 00:01:58.290 --> 00:02:01.800 However, you may be limited to a patient population

38 00:02:01.800 --> 00:02:06.120 from a certain state or medical center.

39 00:02:06.120 --> 00:02:08.970 In our case, we may only have access to patients

40 00:02:08.970 --> 00:02:12.030 with non-alcoholic fatty liver disease

41 00:02:12.030 --> 00:02:17.030 followed in the liver clinic from 2015 to 2020.

42 00:02:17.700 --> 00:02:20.160 In this case, you could either collect data

43 00:02:20.160 --> 00:02:23.610 from all individuals in the available study population

44 00:02:23.610 --> 00:02:26.100 if it's feasible to do that.

45 00:02:26.100 --> 00:02:29.400 Otherwise, if the study population is too large

46 00:02:29.400 --> 00:02:31.710 you could select a random sample

47 00:02:31.710 --> 00:02:34.830 from that available study population.

48 00:02:34.830 --> 00:02:38.550 If you choose a representative random sample

49 00:02:38.550 --> 00:02:42.603 your results are generalizable to that study population.

50 00:02:45.480 --> 00:02:48.600 Next, specify the main intervention,

51 00:02:48.600 --> 00:02:52.320 which is the exposure test treatment

52 00:02:52.320 --> 00:02:54.390 or the main prognostic factor

53 00:02:54.390 --> 00:02:57.000 that you are interested in studying.

54 00:02:57.000 --> 00:03:01.830 For example, lifestyle modification to achieve weight loss

55 00:03:01.830 --> 00:03:04.470 or if studying liver cancer,

56 00:03:04.470 --> 00:03:07.350 your intervention of interest could be serafenib

57 00:03:07.350 --> 00:03:09.033 to prolonged survival.

58 00:03:12.180 --> 00:03:15.330 If you're interested in performing a comparison,

59 00:03:15.330 --> 00:03:18.240 the next step is to specify a control

60 00:03:18.240 --> 00:03:22.500 or comparison intervention or exposure.

61 00:03:22.500 --> 00:03:23.940 This can be, for example,

62 00:03:23.940 --> 00:03:28.533 a placebo control or the current standard of care.

63 00:03:30.240 --> 00:03:33.480 Finally, we must specify the clinical outcome

64 00:03:33.480 --> 00:03:36.510 or primary endpoint of your study.

65 00:03:36.510 --> 00:03:40.380 This includes the element of time, if that's appropriate,

66 00:03:40.380 --> 00:03:42.270 and this would apply if you're looking

67 00:03:42.270 --> 00:03:46.770 at a fixed follow up time period post-intervention.

68 00:03:46.770 --> 00:03:50.790 Say three month survival following surgery

69 00:03:50.790 --> 00:03:54.420 or NAFLD resolution one year following

70 00:03:54.420 --> 00:03:58.353 a certain percentage reduction in total body weight.

71 00:04:00.570 --> 00:04:03.420 Let's run through an example of the type of study

72 00:04:03.420 --> 00:04:07.230 we often perform using medical record data.

73 00:04:07.230 --> 00:04:09.510 The research question asks

74 00:04:09.510 --> 00:04:12.480 among Hepatitis B infected persons,

75 00:04:12.480 --> 00:04:16.560 what factors tests best identify individuals

76 00:04:16.560 --> 00:04:18.930 at highest risk of progression,

77 00:04:18.930 --> 00:04:22.233 as well as those at low risk of progression?

78 00:04:23.370 --> 00:04:27.990 The population studied is Hepatitis B infected persons

79 00:04:27.990 --> 00:04:32.280 treated at the Yale Liver Center between 2011 and 2021.

80 00:04:34.920 --> 00:04:36.540 The interventions of interest

81 00:04:36.540 --> 00:04:39.480 are different patient characteristics.

82 00:04:39.480 --> 00:04:42.870 Specifically, the study will look at different permutations

83 00:04:42.870 --> 00:04:46.740 of key baseline exposures or risk factors

84 00:04:46.740 --> 00:04:51.740 identified in previous studies of Hepatitis B prognosis.

85 00:04:51.780 --> 00:04:54.990 Here, the investigators will look at age

86 00:04:54.990 --> 00:04:58.410 presence of fibrosis, presence of cirrhosis,

87 00:04:58.410 --> 00:05:02.433 elevated ALT, and detectable viral load.

88 00:05:04.590 --> 00:05:07.500 The comparator group for each of these factors

89 00:05:07.500 --> 00:05:10.323 is absence of the baseline factor.

90 00:05:12.120 --> 00:05:15.540 The outcomes of interest are liver related morbidity,

91 00:05:15.540 --> 00:05:17.370 progression of liver disease

92 00:05:17.370 --> 00:05:20.943 and mortality during up to 10 years of follow up.

93 00:05:22.380 --> 00:05:25.170 Now, this is more of an exploratory study

94 00:05:25.170 --> 00:05:28.740 looking for signals of association, but even still,

95 00:05:28.740 --> 00:05:31.440 it has a clearly defined population,

96 00:05:31.440 --> 00:05:34.080 intervention or exposures of interest,

97 00:05:34.080 --> 00:05:37.350 control or reference levels of the exposures

98 00:05:37.350 --> 00:05:39.840 and outcomes of interest.

99 00:05:39.840 --> 00:05:43.320 Sitting down and thinking through the PICO criteria

100 00:05:43.320 --> 00:05:45.480 forces you to make decisions

101 00:05:45.480 --> 00:05:49.443 and pre-specify important aspects of your study.

102 00:05:51.240 --> 00:05:53.160 As we saw in the last example,

103 00:05:53.160 --> 00:05:57.120 there are often multiple clinical endpoints of interest.

104 00:05:57.120 --> 00:06:01.473 Endpoints are classified as clinical or nonclinical.

105 00:06:02.520 --> 00:06:04.740 Clinical endpoints describe outcomes

106 00:06:04.740 --> 00:06:09.540 involving how a patient feels, functions or survives.

107 00:06:09.540 --> 00:06:11.610 They may be assessed by a clinician

108 00:06:11.610 --> 00:06:13.920 and involve clinical judgment,

109 00:06:13.920 --> 00:06:16.293 such as the occurrence of stroke or MI.

110 00:06:17.460 --> 00:06:20.760 They may also be measured by a standard performance measure

111 00:06:20.760 --> 00:06:23.190 such as a pulmonary function test

112 00:06:23.190 --> 00:06:25.170 or they can be patient-reported,

113 00:06:25.170 --> 00:06:28.863 such as self-reported symptoms or quality of life.

114 00:06:30.420 --> 00:06:33.150 Nonclinical endpoints include biomarkers

115 00:06:33.150 --> 00:06:36.660 that may not directly relate to how a patient feels,

116 00:06:36.660 --> 00:06:38.880 however they're thought to be important indicators

117 00:06:38.880 --> 00:06:41.160 of the disease process.

118 00:06:41.160 --> 00:06:44.580 These endpoints can include blood tests, imaging

119 00:06:44.580 --> 00:06:47.883 or other physiological measures such as blood pressure.

120 00:06:48.750 --> 00:06:51.900 A good primary outcome should directly align

121 00:06:51.900 --> 00:06:54.510 with the primary aim of the study.

122 00:06:54.510 --> 00:06:56.850 The endpoint should be accurate

123 00:06:56.850 --> 00:07:01.850 and precise, quantifiable, validated, and reproducible.

124 00:07:02.070 --> 00:07:06.270 We generally include a single primary endpoint.

125 00:07:06.270 --> 00:07:08.700 The goal should be to choose a primary endpoint

126 00:07:08.700 --> 00:07:12.630 that will influence decision making in practice.

127 00:07:12.630 --> 00:07:17.010 The most significant and impactful endpoint that addresses

128 00:07:17.010 --> 00:07:20.940 the research question is chosen as the primary endpoint

129 00:07:20.940 --> 00:07:24.540 and additional important endpoints may be designated

130 00:07:24.540 --> 00:07:26.943 as secondary or tertiary.

131 00:07:28.200 --> 00:07:32.040 Secondary endpoints may not be considered sufficient

132 00:07:32.040 --> 00:07:34.590 to influence decision making alone,

133 00:07:34.590 --> 00:07:37.920 but may help support the claim of efficacy.

134 00:07:37.920 --> 00:07:39.180 Tertiary endpoints

135 00:07:39.180 --> 00:07:42.870 are sometimes called exploratory endpoints.

136 00:07:42.870 --> 00:07:45.360 If included, they are generally used

137 00:07:45.360 --> 00:07:48.063 to test exploratory hypotheses.

138 00:07:49.650 --> 00:07:53.940 Again, we generally use a single primary outcome.

139 00:07:53.940 --> 00:07:56.820 Using multiple primary endpoints may lead

140 00:07:56.820 --> 00:08:01.440 to an unfocused research question and can present problems

141 00:08:01.440 --> 00:08:04.320 with interpretation if the treatment effect is observed

142 00:08:04.320 --> 00:08:07.860 to differ across the multiple outcomes.

143 00:08:07.860 --> 00:08:10.830 However, multiple endpoints may be needed

144 00:08:10.830 --> 00:08:12.750 when a clinical benefit depends

145 00:08:12.750 --> 00:08:15.960 on more than one aspect of the disease.

146 00:08:15.960 --> 00:08:19.620 For example, in Alzheimer's, we may require an effect

147 00:08:19.620 --> 00:08:22.830 on both cognition and function,

148 00:08:22.830 --> 00:08:25.560 so there may be situations where multiple endpoints

149 00:08:25.560 --> 00:08:29.670 are necessary for demonstrating efficacy.

150 00:08:29.670 --> 00:08:32.460 The statistical issue with multiple endpoints

151 00:08:32.460 --> 00:08:35.520 is what we call multiplicity.

152 00:08:35.520 --> 00:08:37.800 When we conduct statistical analysis

153 00:08:37.800 --> 00:08:40.470 and perform hypothesis tests,

154 00:08:40.470 --> 00:08:42.390 there's a chance that we conclude

155 00:08:42.390 --> 00:08:46.530 a significant difference exists between the arms tested

156 00:08:46.530 --> 00:08:49.260 when in truth, there is no difference.

157 00:08:49.260 --> 00:08:52.230 This is due to random variation in the data

158 00:08:52.230 --> 00:08:56.250 that we can observe, but this is a mistake in error,

159 00:08:56.250 --> 00:09:01.250 and we refer to this type of error as a type one error

160 00:09:01.500 --> 00:09:04.530 or the alpha level of the test.

161 00:09:04.530 --> 00:09:07.650 We like to keep this type of error low,

162 00:09:07.650 --> 00:09:12.650 so we typically set the type one error of our tests at 5%.

163 00:09:13.350 --> 00:09:15.660 So when you're testing a single endpoint,

164 00:09:15.660 --> 00:09:19.530 you can maintain a type one error of 5%.

165 00:09:19.530 --> 00:09:23.190 However, suppose we're testing two primary endpoints

166 00:09:23.190 --> 00:09:26.160 and success on either endpoint would lead

167 00:09:26.160 --> 00:09:29.610 to a conclusion of a treatment difference.

168 00:09:29.610 --> 00:09:33.510 The type one error rate on each endpoint compounds

169 00:09:33.510 --> 00:09:36.390 and there's an inflation of the overall type one error

170 00:09:36.390 --> 00:09:39.113 probability above 5%.

171 00:09:39.960 --> 00:09:42.960 This increases the chance of false conclusions

172 00:09:42.960 --> 00:09:46.470 regarding the efficacy of the intervention.

173 00:09:46.470 --> 00:09:48.930 Special statistical testing procedures

174 00:09:48.930 --> 00:09:52.380 need to be used to control the type one error rate

175 00:09:52.380 --> 00:09:54.993 for the study with multiple endpoints.

176 00:09:56.460 --> 00:10:00.150 Multiple primary endpoints occur in three ways.

177 00:10:00.150 --> 00:10:02.610 The first is when there are multiple endpoints

178 00:10:02.610 --> 00:10:04.770 and each endpoint could be sufficient

179 00:10:04.770 --> 00:10:07.230 on its own to establish the efficacy

180 00:10:07.230 --> 00:10:09.510 of the intervention being tested.

181 00:10:09.510 --> 00:10:11.370 These multiple endpoints correspond

182 00:10:11.370 --> 00:10:13.560 to multiple chances of success,

183 00:10:13.560 --> 00:10:16.530 so failure to adjust for multiplicity

184 00:10:16.530 --> 00:10:19.620 can lead to type one error rate inflation

185 00:10:19.620 --> 00:10:23.490 and a false conclusion of effectiveness.

186 00:10:23.490 --> 00:10:27.210 The second option is when the determination of effectiveness

187 00:10:27.210 --> 00:10:30.810 depends on success on all primary endpoints

188 00:10:30.810 --> 00:10:33.960 when there are two or more primary endpoints.

189 00:10:33.960 --> 00:10:37.440 In this setting, there are no multiplicity issues related

190 00:10:37.440 --> 00:10:39.510 to the primary endpoints

191 00:10:39.510 --> 00:10:41.970 as there is only one path that leads

192 00:10:41.970 --> 00:10:45.570 to a successful outcome for the trial and therefore,

193 00:10:45.570 --> 00:10:49.590 no concern with type one error rate inflation.

194 00:10:49.590 --> 00:10:52.320 The third option combines several aspects

195 00:10:52.320 --> 00:10:56.880 of effectiveness into a single primary composite endpoint.

196 00:10:56.880 --> 00:11:01.320 This avoids multiple endpoint related multiplicity issues.

197 00:11:01.320 --> 00:11:03.540 In many cardiovascular studies

198 00:11:03.540 --> 00:11:06.480 it's common to combine several endpoints.

199 00:11:06.480 --> 00:11:11.010 For example, cardiovascular death, heart attack and stroke

200 00:11:11.010 --> 00:11:15.030 into a single composite primary endpoint.

201 00:11:15.030 --> 00:11:17.700 In this case, death is considered on its own

202 00:11:17.700 --> 00:11:19.470 as a secondary endpoint.

203 00:11:19.470 --> 00:11:20.910 If any one of the elements

204 00:11:20.910 --> 00:11:23.400 of the composite outcome is observed,

205 00:11:23.400 --> 00:11:27.090 then the endpoint has occurred for that patient.

206 00:11:27.090 --> 00:11:29.430 It's important that the endpoints included

207 00:11:29.430 --> 00:11:30.960 in the composite endpoint

208 00:11:30.960 --> 00:11:34.320 are of similar clinical importance.

209 00:11:34.320 --> 00:11:37.230 Using a composite endpoint is helpful

210 00:11:37.230 --> 00:11:40.650 when the components are individually rare

211 00:11:40.650 --> 00:11:43.110 so choosing a composite endpoint allows you to

212 00:11:43.110 --> 00:11:45.180 observe more events.

213 00:11:45.180 --> 00:11:48.810 A limitation of using a composite endpoint is that

214 00:11:48.810 --> 00:11:51.090 given the sample size of the study,

215 00:11:51.090 --> 00:11:55.080 there may not be adequate statistical power

216 00:11:55.080 --> 00:11:59.040 to test each component of the endpoint separately.

217 00:11:59.040 --> 00:12:02.400 We'll discuss statistical power in a future video

218 00:12:02.400 --> 00:12:05.160 on elements of sample size calculations.

219 00:12:05.160 --> 00:12:09.360 We'll also discuss endpoints and variables in general,

220 00:12:09.360 --> 00:12:13.593 from a data collection perspective in a future video.

221 00:12:15.090 --> 00:12:17.670 In this video, we discussed important things

222 00:12:17.670 --> 00:12:21.510 to consider when formulating your research question.

223 00:12:21.510 --> 00:12:23.580 From the research question will flow

224 00:12:23.580 --> 00:12:27.570 the specific statistical hypotheses to be tested,

225 00:12:27.570 --> 00:12:30.990 the design of the study, including the sample size,

226 00:12:30.990 --> 00:12:34.560 the data necessary to answer the research question,

227 00:12:34.560 --> 00:12:37.650 the statistical analysis that will be performed

228 00:12:37.650 --> 00:12:40.530 and the conclusions that can be made.

229 00:12:40.530 --> 00:12:44.550 The next video, which is the second video in this series,

230 00:12:44.550 --> 00:12:46.290 will give you an overview

231 00:12:46.290 --> 00:12:50.790 of study designs commonly used in clinical research.

232 00:12:50.790 --> 00:12:54.780 In video three, we will discuss the data collection process

233 00:12:54.780 --> 00:12:58.260 and formally define different variable types.

234 00:12:58.260 --> 00:12:59.850 This video will prepare us

235 00:12:59.850 --> 00:13:03.783 for video four on sample size determination.