Matt Isaac
A01515095
Biostatistics – Homework 6
Due: Monday, 16 April 2018

Exercise 1 (9 points)
In class we discussed five key ethical issues for a RCT. Identify (and comment briefly on) the relevant statistical issues for three of these five ethical issues.

1. Treatment must not be known to be inferior
    a. Statistical issue: This has relation to the 'sidedness' of our test (e.g. <, > or ≠). If we have prior knowledge that one of the treatments is inferior to the other, there may be bias established when we construct our hypothesis.
2. Subjects must provide informed consent
    a. Statistical issue: In statistics, we can only generalize the results of our experiments to populations that were represented in our sample. If participants were coerced into having the treatment, we can only generalize the results to patients who are coerced as well. Thus, we want subjects to voluntarily participate in the study so our results can be generalized to everyone who voluntarily takes the treatment.
3. Sample size should be appropriate
    a. Statistical issue: This is related to the power of our test. We need a large enough sample that we will be able to answer the research question. If we do not have enough participants, we will expose them to risk without being able to answer the research question.

Exercise 2
Psoriasis is a condition involving irritated patches of skin, sometimes to the point of severe flakes or scales. In a RCT considering a new treatment for psoriasis, patients in both treatment and control groups will be evaluated after six weeks by their dermatologist, who will record the proportion of the body covered with scales.

a. (4 points) What type of bias could be present in this RCT, and why?

    Assessment bias could be present. Because the patients' existing dermatologists are recording data, they likely know what treatment/medication the patients are taking. In addition, 'proportion of the body covered with scales' is not a completely objective measurement. The combination of these two factors could lead to biased assessments by the dermatologists.

b. (4 points) How could this type of bias be avoided in this RCT?

    Blinding needs to be incorporated in this study to avoid assessment bias. For example, researchers could have the evaluations done by a dermatologist that has no prior knowledge of the participants, particularly which treatment they have been given.

Exercise 3

Consider a RCT comparing the effects of two treatments, Cerebrolysin and Donepezil, on the cognitive functions of Alzheimer's patients in Spain. Cognitive function is assessed by the ADAS-cog+, a validated, widely used, 14 item psychometric instrument. Patients will be randomly assigned to receive one of the treatments, and their ADAScog+ score will be taken at the beginning of the first week of participation and then 28 weeks later, taking the designated treatment regularly in the meantime. ADAS-cog+ has a maximum score of 85 points with a higher score indicating impairment. A score change of 2 points during the 28 weeks would be considered clinically relevant. The estimated SD for score change during the 28 weeks is 3.3. Researchers want to know if there is any difference in ADAS-cog+ score change (after 28 weeks) between the Cerebrolysin and Donepezil groups.

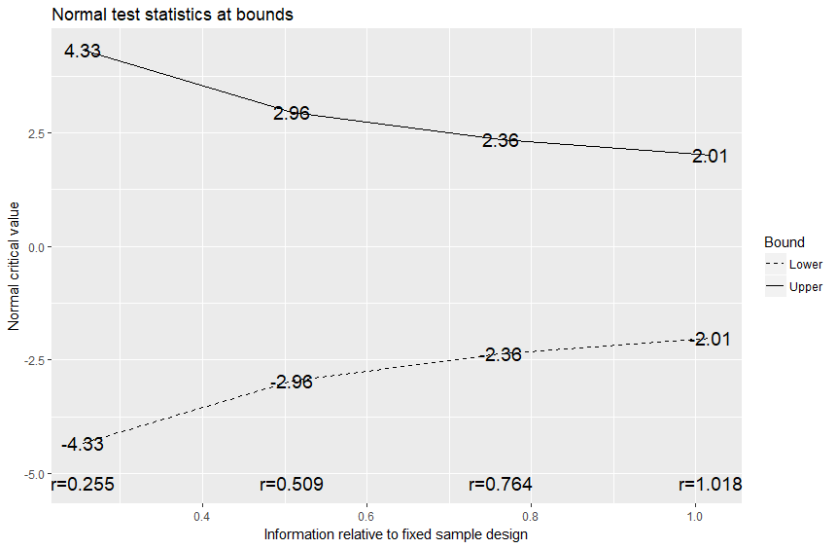
a. (6 points) Assuming an equal number of patients in both the Cerebrolysin and Donepezil groups, how many patients will be required in each group to achieve 80% power in this RCT, testing at level $\alpha = 0.05$?

   We will need at least 44 participants in each group (88 total).

b. (4 points) Rather than conducting a fixed sample size test, why might researchers employ a group sequential design in this scenario?

   It's possible that the more effective treatment will be obviously superior to the less effective treatment. If this is the case, it would save time, resources, and decrease the risk to patients if the superior treatment could be analyzed earlier in the study. Although the researchers might plan on having at least 44 subjects in each group, it is likely that all 44 will not begin and end their 28 weeks of treatment at the same time. Thus interim analyses could be performed that would test for significance. If significant results were obtained, the study could be terminated before subjects are given a treatment known to be inferior.

c. (8 points) Construct and explain a visual display for a group sequential design with 4 total analyses (3 interim, 1 final) for this RCT, using the Lan-DeMets spending function.



The plot above indicates the bounds (upper and lower – indicated by solid and dashed lines) for the test statistic. If the test statistic is outside the bounds at one of the interim analyses, a significant result will have been found and the trial will be stopped. As we get further along in the interim analyses, the interval between the bounds shrinks. This indicates that we are more likely to find a significant result in later interim analyses than at earlier analyses. This is a result of having a larger sample size and therefore more statistical power.

d.

Analysis	Sample Size (per group)	Observed P-value	Significance Threshold
1	0.255 * 44 = 11.220 -> **12**	0.1369	0.0000
2	0.509 * 44 = 22.396 -> **23**	0.03398	0.0030
3	0.764 * 44 = 33.616 -> **34**	0.004184	0.0162
4	1.018 * 44 = 44.792 -> **45**	0.04451	0.0308

The analysis would have been stopped at interim analysis 3.

Exercise 4

Consort Analysis statement:

1 (a)	NA – Title does not indicate that this was a randomized trial.
1 (b)	The experiment was a ’14 week, randomzed, double-blind, placebo controlled study’. With a p-value <0.001, a significant difference in mean change on the depression rating scale was found between the treatment group and placebo group.
2 (a)	

2 (b)	The objective was to investigate the safety and efficacy of aripiprazole as a treatment to depression.
3 (a)	Researchers used parallel assignment and double masking (investigator, and patient).
3 (b)	
4 (a)	In order to included in the study, participants had to be 18-65 years old, and have experienced “single, recurrent, non-psychotic episodes of Major Depressive Disorder, with the current episode of minimally 8 weeks in duration.”
4 (b)	NA
5	The treatment intervention was the administration of the drug Aripiprazole+ ADT. This was administered in oral tablet form. Dose was 2 – 20mg variable dose once daily for 14 weeks. The placebo intervention was the administration of a placebo+ ADT. The placebo was in the same form and dose as the Aripiprazole treatment.
6 (a)	Primary outcome measures: Mean change in the Montgomery Asberg Depression Rating Scale. This is a 10-item, ordinal rating scale. Change was defined as postbaseline score – baseline score.
6 (b)	NA
7 (a)	The primary outcome measure (see 6 (a)) was used to determine sample size. A power calculation was performed with 90% power, assuming standard deviation of 10.5 and two-sided alpha level of 0.05.
7 (b)	NA
8 (a)	No information was given on how randomization was carried out.
8 (b)	NA
9	NA
10	Not listed explicitly, but the study director was Bristol-Myers Squibb
11 (a)	It is indicated that double masking was used. Both the investigator, and patient were blinded from the treatment that was used. How this blinding was carried out is not specified.
11 (b)	NA
12 (a)	A 2 sided t-test was used for all outcomes, both primary and secondary.
12 (b)	NA
13 (a)	See flow chart (next page)
13 (b)	See flow chart (next page)
14 (a)	NA
14 (b)	NA

15

	Treatment	Placebo	Total
Overall	177	172	349
Age (mean, std)	45.1, 10.6	45.6, 11.3	45.4, 10.9
Gender	138 F 39 M	117 F 55 M	255 F 94 M
Ethnicity	Hispanic/Latino: 6 Nonhispanic: 168 Not Reported: 3	Hispanic/Latino: 6 Nonhispanic: 162 Not Reported: 4	Hispanic/Latino: 12 Nonhispanic: 330 Not Reported: 7
Race/Ethnicity	Am. Indian/Alaskan: 0 Asian: 3	Am. Indian/Alaskan: 1 Asian: 2	Am. Indian/Alaskan: 1 Asian: 5

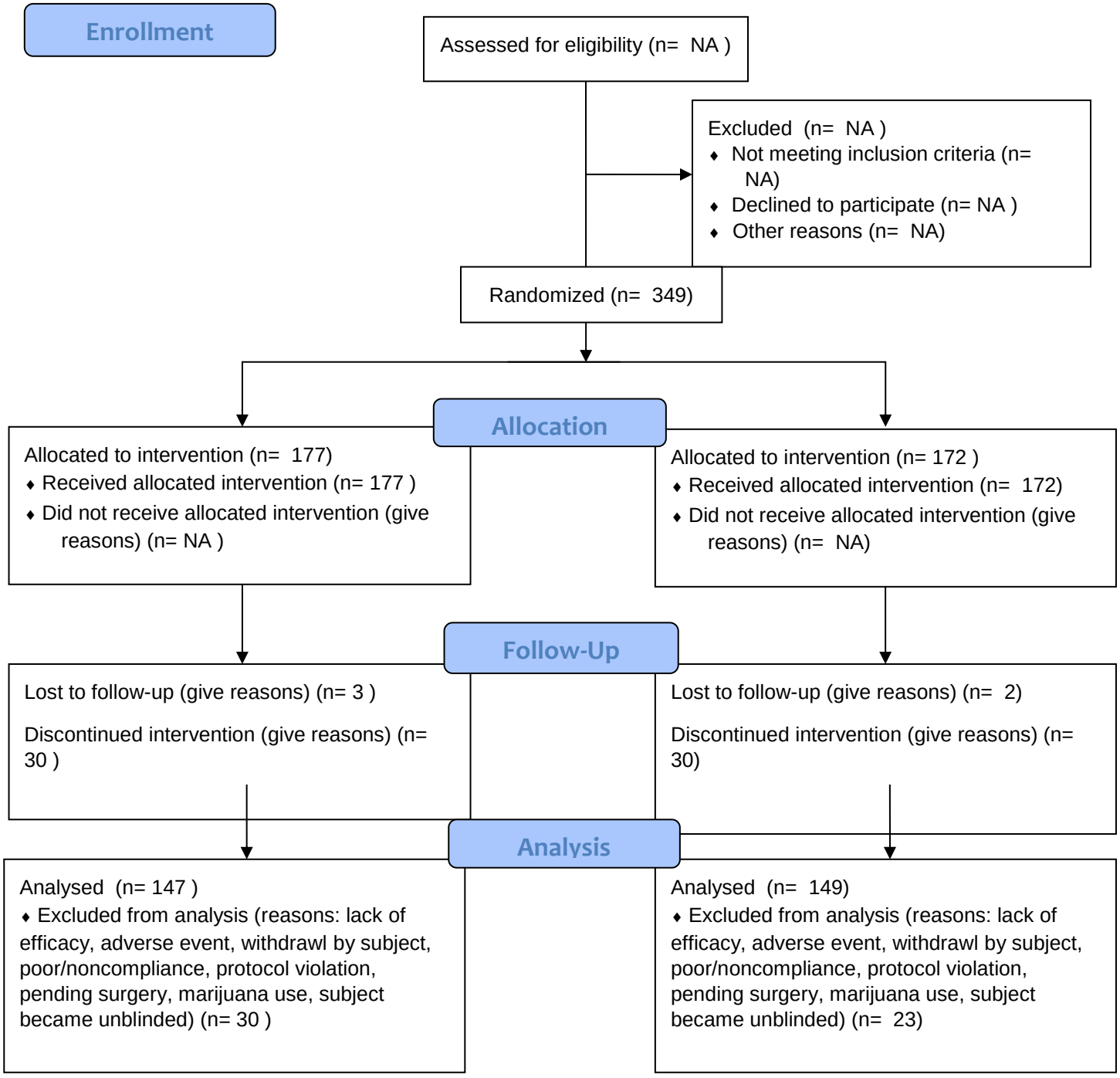
<table>
<tr><td></td><td></td><td>Hawaiian or Pacific Islander: 1
Black: 14
White: 155
Not Reported: 4</td><td>Hawaiian or Pacific Islander: 0
Black: 18
White: 149
Not Reported: 2</td><td>Hawaiian or Pacific Islander: 1
Black: 32
White: 304
Not Reported: 6</td></tr>
<tr><td>16</td><td colspan="4">Number analyzed in each group: Treatment: 174; Placebo: 169</td></tr>
<tr><td>17 (a)</td><td colspan="4">P-value <0.001. 95% confidence interval (-5.44, -2.02).</td></tr>
<tr><td>17 (b)</td><td colspan="4">NA</td></tr>
<tr><td>18</td><td colspan="4">NA</td></tr>
<tr><td>19</td><td colspan="4">Serious Adverse effects: 1 suicidal ideation in the treatment group. 1 arterial occlusive disease in the placebo group.
Other adverse effects:
<table>
<tr><td>Effect</td><td>Aripiprazole</td><td>Placebo</td></tr>
<tr><td>Vision blurred</td><td>13</td><td>3</td></tr>
<tr><td>Nausea</td><td>7</td><td>10</td></tr>
<tr><td>Diarrhea</td><td>10</td><td>13</td></tr>
<tr><td>Constipation</td><td>10</td><td>6</td></tr>
<tr><td>Fatigue</td><td>16</td><td>8</td></tr>
<tr><td>Upper Respiratory Tract Infection</td><td>13</td><td>13</td></tr>
<tr><td>Headache</td><td>15</td><td>14</td></tr>
<tr><td>Akathisia</td><td>32</td><td>6</td></tr>
<tr><td>Dizziness</td><td>9</td><td>5</td></tr>
<tr><td>Somnolence</td><td>10</td><td>1</td></tr>
<tr><td>Insomnia</td><td>15</td><td>9</td></tr>
<tr><td>Restlessness</td><td>22</td><td>6</td></tr>
</table></td></tr>
<tr><td>20</td><td colspan="4">NA</td></tr>
<tr><td>21</td><td colspan="4">NA</td></tr>
<tr><td>22</td><td colspan="4">Nothing explicitly stated, although the p-value from the t-test indicates that there is a highly significant difference in the change of MADRS score between treatment and placebo groups.</td></tr>
<tr><td>23</td><td colspan="4">NCT00105196</td></tr>
<tr><td>24</td><td colspan="4">NA</td></tr>
<tr><td>25</td><td colspan="4">This study was sponsored by (and collaborated with) Otsuka Pharmaceutical Development & Commercialization, Inc. and Otsuka America Pharmaceutical.</td></tr>
</table>

NOTE: CONSORT statement is below, with R code below that.

CONSORT

TRANSPARENT REPORTING of TRIALS

CONSORT 2010 Flow Diagram



Code Appendix:

```
 # Exercise 3 (a)
power.t.test(sd = 3.3, sig.level = 0.05, power = 0.8, delta = 2, type = "two.sample")


# Exercise 3 (c)
library(gsDesign)
d <- gsDesign(k = 4, test.type = 2, alpha = 0.025, sfu = sfLDOF)
plot(d)

# Exercise 3 (d)
rct <-  read.csv("http://www.stat.usu.edu/jrstevens/biostat/data/RCT.csv")
hist(rct$scorechange[which(rct$trt == "Cerebrolysin")])
hist(rct$scorechange[which(rct$trt == "Donepezil")])

# Assumption of approximate normality met. T test will be used.
t.test(rct$scorechange[1:12], rct$scorechange[101:112], paired = FALSE)
t.test(rct$scorechange[1:23], rct$scorechange[101:123], paired = FALSE)
t.test(rct$scorechange[1:34], rct$scorechange[101:134], paired = FALSE)
t.test(rct$scorechange[1:45], rct$scorechange[101:145], paired = FALSE)
```