Week 4 Experimental Design

SIADS 688 Data Science for Social Good

Section 1. Introduction (3 points)

Write a 200-word introduction that defines the research topic and answers the question, "Why does this topic matter?" Read the Introduction section of the relevant papers we have covered to learn how to write an effective introduction. In this section, you should do the following:

1. **(1 point) Establish the problem leading to the study. Please narrow your topic as much as you can.**
2. **(1 point) At some point in your discussion, identify the questions you will try to answer through your research. Make these as concrete and focused as possible, so that your reader can figure out exactly what you want to learn.**
3. **(1 point) The question "Why does this topic matter?" implies the sub-question "To whom does this matter?" What's your audience, and why should they care?**

The ability to influence behavior via social information and intervention has been well-documented. There have been studies that attempt to influence driving behavior via social comparison, for example, as individuals take into consideration the behavior of those like oneself and of those perceived to be of a higher status. Despite the prevalence of these types of studies, there are still multiple fields that warrant further investigation to determine how social information can impact behavior. One area is the health care space, specifically for older individuals with heart failure. Heart failure is a major cause of hospitalization in the United States for the elderly population. Not only is the disease debilitating in terms of impact on day-to-day life and activity levels, but heart failure also can stretch the effectiveness of hospital systems. Individuals with heart failure commonly enter through the emergency department when acute flareups occur, which occupies beds and personnel that could be served for other types of emergency care. Alleviating this issue would be a social good, as not only would elderly health improve, but hospital resource strain would decrease. With heart failure admissions being such a prevalent issue within the healthcare space, there are many questions around how to best address this issue. In particular, can social information be provided via text or email to direct an elderly population towards a lifestyle that can reduce the risk of heart failure? The goal of this study would tailor to healthcare points of contact who'd be able to provide social information to their patients to assist them in leading a healthy active lifestyle to help reduce the risks of this debilitating disease.

Section 2. Literature Review (3 points)

Write a 300-word literature review for your chosen topic that answers the question, "What do we already know about this topic, and what do we still need to find out?" The purpose of the literature review is to understand how the new knowledge you will be acquiring fits into what is already known. You need to discuss at least 3 items from the literature. You can start with the classic papers in an area of research, then do a Google Scholar search to find the most relevant papers.

Heart failure precedence in the United States' elderly population is significant and places a notable strain on the United States' healthcare infrastructure. Heart failure, while there are different variations, commonly impacts day-to-day life through symptomatic shortness of breath, chest pain, etc. With these types of symptoms, it's very common for individuals suffering from this illness to enter the hospital system through the emergency department, consuming large amounts of care and hospital resources to stabilize. As such, this is a common research area in the healthcare space in cardiology. There are numerous clinical trials that focus on individuals already with the illness and how to best stabilize them when they're out of the hospital. However, the specific area of interest for this work is how to promote a healthier lifestyle to avoid succumbing to this illness to begin with. Understandably, and perhaps unsurprisingly, doctors do have recommendations for their patients around

how to best avoid the risk of heart failure and other cardiovascular diseases. These recommendations include a consistent workout regimen with consistent physical activity, eating healthy foods, weight management (which is tied closely to the previous two recommendations), and, lastly, avoiding smoking, with the pulmonary and cardiovascular systems so closely intertwined. While these best practices due notably reduce the risk of heart failure and other cardiovascular diseases, it is often a challenging task to either motivate or inform patients to follow these types of guidelines. Research into social intervention via text or email messaging is still in its infancy. However, there have been successful text message interventions related to diabetes management and medication management when dealing with viral diseases. The goal of this study will be to understand the effectiveness of text message social intervention in changing elderly behavior to reduce smoking and improve day-to-day physical activity and dieting. One particular concern related to this study design would be the elderly and their often-limited use of technology due to lack of experience, comfort, etc. with these tools. However, promising studies have emerged recently that conclude this population is willing and likely to use technology following exposure and training. Therefore, a text messaging intervention to improve health outcomes will be the focus of our study.

Section 3. Experiment Design (10 points)

Hypotheses (2 points)

for each hypothesis, state both the null and the alternative hypothesis. For each hypothesis, state where it comes from, either theory or past empirical or experimental findings. Cite relevant papers which support each hypothesis.

Hypothesis 1

Null Hypothesis: Providing a weekly text message on the benefits of increased activity levels does not increase daily physical activity and, therefore, does not help reduce the risk of heart failure.

Alternate Hypothesis: Providing a weekly text message on the benefits of increased activity levels increases daily physical activity, an important factor in reducing the risk of heart failure.

This hypothesis is of interest as past empirical studies have made clear the importance of high levels of physical activity in the elderly population to help reduce the risks of heart failure and other cardiovascular disease. While there have been studies that have leveraged text messaging to intervene in diabetes and medication management, this is a novel study assessing the ability to socially intervene and correct behavior leading to better health outcomes and reduced hospital strain.

Hypothesis 2

Null Hypothesis: Providing a weekly text message on the benefits of quitting smoking does not change smoking pattern or behavior, and, therefore, does not help reduce the risk of heart failure.

Alternate Hypothesis: Providing a weekly text message on the benefits of quitting smoking does help reduce smoking, an important factor in reducing the risk of heart failure.

This hypothesis is of interest as past empirical studies have made clear the importance of quitting smoking to help reduce the risks of heart failure and other cardiovascular disease. While there have been studies that have leveraged text messaging to intervene in diabetes and medication management, this is a novel study assessing the ability to socially intervene and correct behavior leading to better health outcomes and reduced hospital strain.

Experimental Procedure (5 points)

The procedure should include subject recruitment, software used, and experimental variations. You should answer the following questions:

a) **Who are your subjects? For example, they can be college students, MTurkers, or users of a website.**
b) **What are the treatment and control conditions? You may have multiple treatment groups, and control groups.**
c) **What are the outcomes you would like to measure and why are these outcome variables relevant? Interpret the outcomes you plan to measure.**
d) **What are the potential confounders in your experimental design?**
e) **Is your design a factorial design, block design, cluster design, or none of the above?**

The subjects of this study will be an older population of 40–50-year-old individuals, which is younger than but not far off from the age range where heart failure begins to become particularly prevalent. To avoid any types of bias related to race and gender, the study should ensure a diverse subject pool. For the first hypothesis, all 40-50-year-old individuals can be considered, whereas for the second hypothesis, this study would only focus on a subsection of the population that identifies as smokers.

In terms of the treatment and control conditions, for the first hypothesis, the subjects in the control condition will receive a generic text message from their doctor's office around scheduling appointments to ensure they're up to date on care. For the treatment condition, the subjects will again receive a text message but this time encouraging 30 minutes of exercise a day with additional information included in the message noting the benefits of increased physical activity for heart health. Similarly, for the second hypothesis, the subjects in the control condition will receive the same text message as the control group in the first hypothesis. For the treatment condition, the text message will encourage smoking cessation while providing additional information noting the benefits of smoking cessation for heart health.

In terms of the outcomes that are to be measured as part of this study, we would assess over an arbitrary period of 3-months the number of individuals that increased their daily physical activity (for hypothesis 1) and the number of identified smokers who quit smoking during the 3-month period (hypothesis 2). In terms of assessing physical activity, there have been heart failure studies already that have tracked this information, leveraging belt accelerometers to track daily footsteps. These outcome variables are relevant as increased physical activity and smoking cessation are positive health improvements that would not only decrease the risk of developing heart failure in the future, but, as a result, would be for the social good in helping alleviate healthcare resources dedicated to treating this debilitating illness. For hypothesis 1, the outcome that we would measure would be increased activity as measured by the belt accelerometer. For hypothesis 2, smoking cessation would be the desired outcome, and we would assess the percent of subjects that ceased smoking in the treatment group relative to the control group.

In terms of confounders, technology-savvy is a factor that will need to be carefully considered in the subject pool. There may be cases where a more elderly population has individuals that are less familiar with technology and may not be as receptive to receiving these types of updates. Additionally, the time of year may be an interesting confounder to focus on. In winter months, it may be reasonable and expected for physical activity to increase and smoking to perhaps even increase due to worsening weather and the impacts this can have on mental health. Whereas in summer months, physical activity may naturally increase due to the warmer weather. It will be important to consider the seasonal impacts, therefore, that may impact physical activity and smoking cessation studies.

This will be a block design study with subjects being randomly assigned into treatment and control groups to address both hypotheses.

Sample size calculation using power analysis (3 points)

You can use (1) data from previous studies, e.g., their σ and δ; or
(2) the σ/ δ ratio to compute your sample size. Note that you cannot take part of your σ and δ from a previous empirical study and force a σ/ δ ratio. If you use empirical data, you need to use their σ and δ in a consistent way.

For this sample size calculation for both hypotheses, we will rely on standards where alpha = 0.05 and power is 0.80 (beta being 0.20). In heart failure related studies, most researchers also focused on detecting differences of 0.5, or ½, standard deviations, and therefore, the sigma/delta ratio for this study will be 2.

With our alpha, we know that 1 – alpha/2 is 0.975, which has a z-score of 1.96. With our beta, we know that power is 0.80, which would have a z-score of 0.84. Using this information, we can calculate our sample size as:

$2(1.96 + 0.84)^2 * (2)^2 = 64$

Therefore, we should have 128 subjects total when evaluating Hypothesis 1 (64 in both the treatment and control groups) and 128 subjects total when evaluating Hypothesis 2 (64 in both the treatment and control groups).

Section 4. Conclusions and Limitations (1 point)

This experiment will examine social intervention in the healthcare landscape. While there has been some success social engineering appropriate adherence to diabetes and other medications, this would be the first such study that would leverage social intervention in the form of text messaging to alter behaviors to achieve better health outcomes for patients and reduce hospital strain. Through text messaging, healthcare providers would hopefully be able to interact with their patients in a more novel way and achieve a desirable result, a patient less likely to undergo heart failure due to smoking cessation and increased physical activity.

In terms of limitations, as noted previously, a major confounding variable around physical activity and smoking cessation would be the time of year such a study was run and the impact that weather and the seasons would have on behavior. Overall, with a long enough window to evaluate the impact of this potential confounder, this study is still worthwhile to better understand the impact social intervention may have on driving meaningful healthcare outcomes.

Appendix: Experimental Instructions or Materials (2 points)

This appendix should contain the script that you read or send to your subjects for your experiment, one for each experimental condition. a. Lab experiments: This will be the instructions you will read to the subjects. It should include review questions to test the subjects' understanding. b. Field experiments: This might be the emails or other stimuli you send to the participants to implement your treatment(s) and control.

<u>Control Text Message in both the Hypothesis 1 and 2 Studies:</u>

Hi [subject name] it's time for your yearly cardiology appointment. Please call us today to schedule a time for your visit.

<u>Treatment Text Message for Hypothesis 1 Study (Physical Activity):</u>

Hi [subject name] it's time for your yearly cardiology appointment. Please call us today to schedule a time for your visit. Ahead of your visit, a reminder that even just 30 minutes of exercise a day has notable impacts on heart health and reduces the risks of developing various cardiovascular diseases including heart failure.

Treatment Text Message for Hypothesis 2 Study (Smoking Cessation):

Hi [subject name] it's time for your yearly cardiology appointment. Please call us today to schedule a time for your visit. Ahead of your visit, a reminder that quitting smoking has notable impacts on heart health and reduces the risks of developing various cardiovascular diseases including heart failure.

References (1 point)

This section should contain at least the three references used in your literature review.

Hall AK, Cole-Lewis H, Bernhardt JM. Mobile text messaging for health: a systematic review of reviews. Annu Rev Public Health. 2015 Mar 18;36:393-415. doi: 10.1146/annurev-publhealth-031914-122855. PMID: 25785892; PMCID: PMC4406229.

Margulies KB, Hernandez AF, Redfield MM, et al. Effects of Liraglutide on Clinical Stability Among Patients With Advanced Heart Failure and Reduced Ejection Fraction: A Randomized Clinical Trial. *JAMA.* 2016;316(5):500–508. doi:10.1001/jama.2016.10260

Redfield, Margaret M., et al. Isosorbide Mononitrate in Heart Failure with Preserved Ejection Fraction 2015/11/08. New England Journal of Medicine 2314-2324. Volume 373. 24 10.1056/NEJMoa1510774 26549714.

Rippe JM. Lifestyle Strategies for Risk Factor Reduction, Prevention, and Treatment of Cardiovascular Disease. Am J Lifestyle Med. 2018 Dec 2;13(2):204-212. doi: 10.1177/1559827618812395. PMID: 30800027; PMCID: PMC6378495.

Vaportzis E, Clausen MG, Gow AJ. Older Adults Perceptions of Technology and Barriers to Interacting with Tablet Computers: A Focus Group Study. Front Psychol. 2017 Oct 4;8:1687. doi: 10.3389/fpsyg.2017.01687. PMID: 29071004; PMCID: PMC5649151.