Meeting Minutes 21 Aug 2020

Link: https://www.notion.so/xmeng/Meeting-21-Aug-2020-41b64f0245ea440a8557e97f8cc11286

People present

Susan, Peng, Walter, Xiang

Thoughts on pre-processing the data

Start / End day

Walter: Set the start day to one when people start to receive messages reasonably.

Susan: Could set to **1 Sep 2019—1 Sep 2020**. *Trim the data* if the people's follow up time is outside of this range.

Baseline feature

Susan: Add a **binary variable** indicating whether the person joined before 1 Sep 2019 and how long they have been under treatment prior to 1 sept 2019. Also I recall that each person has a week of just data collection with fitbit tracker, then a week with MRT with activity suggestions (in this second week are people randomized to anti-sedentary messages as well? ask Peng) and then they "start" experiencing the bandit algorithm for the activity suggestions. So baseline data can include data from first or maybe even second weeks.

About the study

0. HeartStep has 3 versions: V2; V3; U (very little data);

Some implementation details (Worth checking again either on data or on the document)

1. Average is 1.5 anti-sedentary messages per person per day.
2. Sedentary = less than 150 steps in the past 40 minutes

Outcome brainstorming:

Change in Heart rate

- It's a direct measure of whether people standup, which is the goal of this particular treatment component

Problems:

> 💡 1. Heart rate due to the same exercise can **vary** among different people

Need a way to operationalize what it means "change in heart rate"

Also can we make use of the 3D location data (I cannot spell that terminology)

> 💡 2. What variable? Do we have access to the gyroscope and accelerometer data (Please ask Nick. I do think we do and I am not sure we what this data.... We'd have to have a clear plan)

- Indicator(Mean of x min HR after the message > mean of x min before the message)? What should x be if we choose this?
- Standardized(Mean of 5 min HR after the message - mean of 5 min before the message)? If we want a continuous outcome then we need to find ways to **standardize the outcome**. Eg: Regress on weight, since the change of heart rate might be significant for people in a standard shape, and vice versa.

Heart rate & Steps

Not recommended because people could stretching their arms and do not walk

> 💡 Xiang's new question: If they stretch their arms, do not walk, and their HR increases, we achieve out goal, but the algorithm still think the person is sedentary/risky 5 mins later right?

Susan's reply: Yes this is try but availability considerations prevent the app from sending another anti-sedentary message —information for availability is documented somewhere—ask Peng to tell you where.

Latent Variable Model

Might go to another direction rather than our aim

Things to do during next week (24 Aug-30 Aug)

1. Some Initial EDA:
   - [ ] Histogram for individual: #Messages vs Days
   - [ ] line chart showing a bunch of people's messages vs Days

Xiang's personal note

1. Connect with Pedja and Nick. Get access to

   1) github site

   2) google buckets that store service data and nightly data dump

   3) dashboard page

2. Problems to check / Questions to ask

   What are obvious bugs and potential bugs?

   Which part did the algorithm go wrong?