Analytical Exercise

Whether or not we should focus on a small number of clinics or larger number of clinics largely depends on the *nature* of the experiment. Here the question is how receiving a text message reminder effects vaccination uptake. Another point to note is that the cost for the treatment and control is *not* the same. It costs $1 to send a text message, so providing the treatment is more expensive. Another aspect to be vary of is that there is a cost of $100 for each clinic we would like to participate with.

While conducting such a study we must also keep in mind SUTVA, which is the Stable Unit Treatment Value Assumption. It takes into consideration the possibilities of spillovers and relevant treatment - in this case, it is whether sending a text message to the treatment might invariably have an effect on the control group, and also whether sending a text message might trigger *other* effects which might have an effect on the control group.

Since we are limited in the number of clinics we can choose by the fixed cost, it would seem that we should choose a smaller number of clinics; we assume that spillover is fairly well contained between multiple clinics here because text messages are not public and will only effect the person you send it to. For example, if we choose 2 clinics, one which acts as the control clinic, and the other where we send text messages to all participants, we have spent $800, and can use difference in means to evaluate if there has been an effect. Here there are an equal number of treatment and control, and by choosing different clinics we further control the spillover effect.

If we further feel that there is no spillover effect at all, we can afford to choose multiple clinics. In such a case, maybe choosing 4 clinics, and 150 patients in each clinic would also lead to a sample which can effectively identify the effect of text messages. One of the main considerations we have to keep in mind when choosing the number of clinics is the *kind* of spillover effects (between clinics, within the same clinic). Based on what we can control better we should split up the clinics accordingly.

A spillover effect in the *same* clinic suggests choosing only 2 clinics, one treatment and one control. This way even if there is a spillover effect within patients in one clinic, they all still receive the same treatment. If there is not much of a spillover effect then we can afford to spread out our treatment across clinics. Spreading out the treatment across multiple clinics ensures that we are not measuring any hidden treatments, for example if a clinic is in a particular neighbourhood where people vaccinate more.

The above factors would also effect the smallest effect size we will be able to reliably detect. We need to carefully consider the possibility of spillovers within and between clinics as well as the possibility of hidden treatment based on where a clinic might be located, or the demography of the people attending a particular clinic.

Using a smaller number of clinics would suggest we use a difference in means method of measuring effect size, and if we choose multiple clinics we should consider using difference of differences because it is a mixed setup. Since the smallest effect size depends on the kind of participants we choose from the clinics, and because the kind of participants we choose depends on spillover effects and hidden treatments, these are the factors which determine the effect size.