**MSCR 520: Designing Clinical Trials**

**Final Examination**

**22 April 2020**

**INSTRUCTIONS: Please complete legibly by printing if necessary. If your response is not legible, credit will not be given. Please note that the Honor Code for the Graduate School of Arts and Sciences is in effect for the duration of the examination. You may not consult any resource or person while completing this examination. Please write your name on EACH page. Please make sure that you have 6 pages.**

**Maximum time: 120 minutes**

**Final Score:\_\_\_\_\_\_\_\_\_**

**POINT DISTRIBUTION: Indicated next to each question. Partial credit will be given, so please show your work/ thought processes. Total points: 60**

**Thank you and good knowledge!**

**-----------------------------------------------------------------------------------------------------------------**

1. **For each of the 5 following study designs**, briefly describe (define) the design and list one advantage and one disadvantage for each of the selected study designs: (3 points each for total of 15 points)
2. Randomized controlled trial:

1) Define: a comparative study with an intervention group and a control group, assigned by randomization.

2) Advantage: you can control for known and unknown variables, and remove potential for bias

3) Disadvantage: overall expensive, and time consuming. There has also been ethical debates of potentially depriving participants from receiving an intervention or treatment.

b. Cross-over design:

1) Define: it is a kind of RCT, this design allows for study participants to be their own “control”, because there are 2 periods during study. In the first period a participant will receive control or intervention, and in the second period the alternative

2) Advantage: Researchers can assess how participant does on both control and intervention, and variability is reduced. By doing this you can use a smaller sample N to find a difference.

3) Disadvantage: Need for washout period between study phases, so that the effect of intervention in first stage is not carried into the latter.

c. Observational design:

1) Define: as the name implies, investigators observe study participants without an intervention, in contrast to RCT where one intervenes and looks for the effect of such intervention on an outcome.

2) Advantage: You can establish correlation. Systematic and unbiased observation can result in a true picture of individual’s natural set of behaviors.

3) Disadvantage: Cannot establish causation. Can be time-consuming depending on outcome. Researcher bias

d. Factorial design:

1) Define: In a single experiment, researchers attempt to assess 2 interventions compared to control.

2) Advantage: More efficient that one-factor-at-a-time experiments due to smaller sample size. Less costly.

3) Disadvantage: An important assumption is that there no interaction between treatments under investigation. One must consider potential adverse for polypharmacy, and could lead to less adherence from participants.

e. Bayesian inference:

1) Define: is a method of statistical **inference** in which **Bayes**' theorem is used to update the probability for a hypothesis as more evidence accumulate.

2) Advantage: More information for decision making. Sample size reduction via prior information or via Adaptive Trial Design. Often easier to implement using Bayesian than frequentist methods.

3) Disadvantage: Extensive preplanning, model building, computationally challenging and choice of prior information may need to be justified.

1. **In deciding whether a subject receives the intervention or the placebo, why is it important to use a randomization procedure? We learned simple, blocked, and stratified randomization. What is the advantage of blocked randomization over simple randomization? (4 points)**

By randomization, we limit the bias (conscious or subconscious bias) in the allocation of patients to either control or intervention groups. Also by randomizing, we are distributing all factors (both known or unknown) by chance and not patient preference.

In block randomization, the order of the interventions that would be assigned in each block are randomized. This way we guarantee that at no time during randomization, there would be an imbalance in the number of participants assigned to either controls or intervention group. This is a major advantage to simple method where there could always be a possibility of substantial imbalance at any point in the randomization.

1. **A clinical dietician wants to compare two different diets, A and B, for diabetic patients. She hypothesizes that the mean blood glucose level of subjects given diet A will be lower than the mean blood glucose level of subjects given diet B. The investigator plans to take a random sample of diabetic patients and randomly assign them to one of the two diets. Fasting blood glucose will be measured at baseline and at the end of 6 weeks. The primary outcome is the change in fasting blood glucose levels at the end of 6 weeks.**
2. What information does the dietician need to calculate the sample size? (4 points)

Standard deviation of BG

Mean BG of groups (or difference in means that is clinically important)

Alpha

Power

1. How do the different pieces of information affect the sample size value? (4 points)

Alpha = smaller the alpha, larger sample size will be needed

Power = bigger the power, larger sample size needed

Difference in means = the smaller the difference, the larger the sample size needed

SD = larger the SD, the larger the sample size needed

**4. In the previous problem, the following regression model is considered at the analysis stage:**

**Change in Blood Glucose = β0 + β1 DIET\_A + β2 FEMALE + ε**

**where DIET\_A is 1 for subjects who are given Diet A and 0 for subjects who are given Diet B. Moreover, FEMALE is 1 for female subjects, 0 for male subjects. Interpret parameters β1 and β2. (4 points)**

B1 is the estimated unit change in blood glucose that is expected to occur with having diet A instead of diet B, controlling for the effect of sex.

B2 is the estimated unit change in blood glucose that is expected to occur for being female compared to male, with adjustment/control for diet group.

5. A new study team member tells you that he shredded a pile of papers while organizing in preparation for a site visit last week. You realize that the team member shredded paper case report forms for 5 active subjects enrolled in your industry-sponsored clinical trial. Fill in the blanks describing the problem, correction, risk of severity, and reporting (4 points total, 1 point each)

1) Problem: Complete disregard for good clinical practice, possible data loss and breach of data.

2) Correction:

– Evaluate rights, welfare, and safety of subject

– Report, if applicable: subject, subject’s family,

sponsor, and/or IRB

Conduct root cause analysis to identify potential contributing factors and report the issue to the appropriate. Verify and conduct CAPA to ensure no recurrence of similar ensure in the future.

3) Risk of severity to subjects: No Minimal- there is possible risk of having side effects to the intervention and we could not verify the presence or lack of intervention.

4) Reporting (to whom): Data safety monitoring body and Sponsor.

1. **You are conducting a clinical trial in which you hope to assess the immune response of healthy infants who receive a probiotic compared with those infants who receive a placebo. You are having difficulty with recruitment because many of the parents of potential subjects do not want their child to participate as there is a chance that they might receive a placebo.**

**Describe a strategy identifying a study design that you could implement to overcome this study recruitment challenge (1 point) and how this study design would overcome the challenge of being assigned to receive a placebo.**

**a. Study design: (1 point)**

Historical controls

Cross-over

Interrupted times series trial

**b. How does the study design described in your response to 6a overcome the perceived challenge of assignment to the placebo group? (1 point)**

The interrupted time series allows us to use patients over time to study the “slope” of their immune response, adjusting for any time-dependent variables, and then adding the intervention of probiotics to all patients. Then, after the intervention point you can study the immune response pattern over time. If the new “slope” is significantly different, we can infer an effect of the intervention.

This way we will use historical clinical trial data on the control arm, so more trial resources can be devoted to the active treatment arm. With this method we will have more accurate point estimates, increased power, and reduced type I error

1. **a. How does the FDA define an Adverse Event? (1 point)**

Adverse Event = untoward medical occurrence associated with use of a drug in humans (whether or not it is drug related)

b**. What is a CAPA plan? (1 point)**

CAPA = corrective and preventative actions; corrective actions is process of reacting to existing problem, while preventive actions are process of detecting potential problems and eliminating them

c**. What does GCP stand for with respect to clinical research (1 point) and what is its relevance? (1 point)?**

GCP (good clinical practice): national ethical and scientific quality standard for designing, conducting, and recording and reporting clinical trials

Relevance of GCP to clinical research: GCP training is requiring for all investigators in FDA-regulated clinical trials. It involved IRB approval, protocol adherence, safety recording/reporting, data and safety monitoring, quality assurance/improvement

1. Multiple imputation is generally favored over single imputation or complete case analysis. Explain why. (4 points)

Multiple imputation is better as it is flexible and allows uncertainty due to missing values while reducing the probability of random error as compared to single imputation.

Missing data, when ignored or replaced with just a median, can lead to biased estimates, wrong standard errors, and invalid statistical inference. Data may not be missing at random, or it may be related it its value or the value of other data. Single imputation assumes MCAR, which is rarely true, and uses distributions that have too many cases at the mean (and thus parameter estimates are biased towards zero), and narrower confidence intervals (thus more type 1 error).

Multiple imputation:

* Incomplete multivariate data is filled with m cases to generate m datasets.
* M datasets are analyzed with regression procedures
* M data sets are pooled or combined before interpretation.

1. O’Brien-Fleming is a popular method used in planning for interim analyses. How does the O’Brien-Fleming method address the problem of repeated testing for significance? (4 points)

Repeated testing increases the probability of type 1 error. However, O’Brien-Fleming model lead to lower likelihood of stopping in the early stages. However, this procedure leads to a greater chance of stopping prior to the end of the study than Pocock.

O’Brien-Fleming: most common. It uses Z test statistic, with Z\* = sqrt(K/i). Essentially, each interim analysis has a decreasing critical value until your final test. Your final test will essentially be 1.96 (similar to what would be expected if only single test was used). However, requires greater conservatism with early analysis.

1. What is a prior distribution, and what are some sources of prior information? (4 points)

In Bayesian analysis, the posterior probability is proportional to likelihood x the prior probability. The prior probability depends on the prior distribution (the range of values that are believe to be known at some level of confidence). Prior distribution can help combine with study information to create and incorporate results into the prior information. Prior information also reduces the need for a larger trial (e.g. adaptive trial design). Prior information is either informative (gives values as being more likely, thus skewed) or uninformative (a uniform distribution). Sources for prior information include clinical trials, registries, pilot studies, or expert opinions. However, the less evidence or data available will decrease the weight of prior information.

1. Name and describe a responsibility of four different members of a clinical trials team? (1 point each/4 points total)
   1. Principal investigator: team leader, supervises all aspects of a clinical trial. Develops the concept and study protocol. Responsible for team, and submits to IRB. Supervises data collection, interpretation and presentation
   2. Data manager: manually enters data collected, works closely with PI to identify the data needed, prepares summary of data for analysis
   3. Research nurse: educates participants, referring physicians about the trial. Communicates regularly with PI and study participants. Assists with informed consent, data management.
   4. Biostatistician: processes and analyses data collected.
   5. Research coordinator: organizes study staff into assignments/tasks, organizes meeting times, will help coordinate with the study staff and participants
   6. Recruiter: responsible for helping with identifying potential study patients, as well as recruiting them into the study (making study materials, media, and may also help with consent)
   7. Pharmacist
   8. Lab Technician

12. a. **Name your favorite study design discussed in class and describe which attributes make this your favorite study design (1 point):**

Point of care clinical trial. For me this was the most interesting trial, because I am working in the VA system with a long history of access to EMR (CPRS system). Using this type of clinical trial, I would be able to more rapidly do a clinical trial which would be less costly an be able to apply those results in my clinical setting immediately.

Interrupted time series analysis:

Group sequential design:

**b. What is a 1572 (1 point)?**

Form 1572 (statement of investigator) is an agreement that is signed by each investigator to provide information to the sponsor, ensuring compliance with FDA regulations for a clinical trial

c. What clinical trial design element should be reduced when trying to establish scientific rigor? (1 point)?

–Not addressing statistical power

–P-hacking, HARKing, fishing expeditions

–Using poorly defined/unvalidated outcome measures