*Serologic Response to Vaccine for COVID-19 in Patients with Hematologic Malignancy*

C. W. Phua, Chin-Yee, B. Hedley, V. Bhayana, S. Lam, S. Li, J. Delport, M. Knauer, M. Haeryfar, S. Shalhoub, J. Kim, A. Quint

Methods: Statistical Analysis

Table of Contents

[Primary Objective 3](#_Toc77775238)

[Secondary Objectives 3](#_Toc77775239)

[*Anticipated Results* 5](#_Toc77775240)

[Sample size determination 6](#_Toc77775241)

[Additional Questions/Concerns 7](#_Toc77775242)

[Sources: 7](#_Toc77775243)

# 

# **Primary Objective**

1. To evaluate the overall seroconversion to the COVID-19 vaccination in patients with hematologic malignancies. The serologic response after each vaccine dose will also be evaluated using the established Elisa antibody testing.

Notes:

* Question: The excel spreadsheets don’t have any data or columns referring to lab values (i.e., antibody levels, etc.). No column on comorbidities either, but I think we decided to exclude this.
* Answer: Lab results will be placed in later. We’ll ignore comorbidities
* Question: How are we assessing vaccine efficacy at the different time points? Are we checking for antibody concentration (continuous variable), whether antibodies are present (Y/N, i.e., categorical), or something else? Based on the Roche Elecsys Anti‑SARS‑CoV‑2 immunoassay, antibody concentration would be best.
* Answer: Vaccine efficacy is based on who gets covid after getting vaccinated.
* Question: What is our control/comparator group for the serologic data? (For ex., comparing antibody concentrations at each time point for people with hem. Malignancy vs. without)
* Answer: Won’t have control group. Could explore data from outside literature.
* .

Statistical tests:

Use repeated measure ANOVA to compare the means of the antibody concentrations (assuming continuous variable) at the different time points. Can do 2 repeated measure ANOVA’s, one for the first timeline, and the next for the second timeline. Can do another repeated measure ANOVA using the entire dataset, but effectively excluding the “Within 4 weeks of dose 2” timepoint.

If we have data on the “normal” (i.e., healthy individual) antibody concentrations from a different population of people at those given timepoints, can do a t-test for each time point to determine if antibody concentrations for those with hematologic malignancies are significantly lower.

# **Secondary Objectives**

1. To explore the variables that may affect antibody response such as type of hematologic malignancy, treatment received in proximity to vaccination, innate level of immunity pre-vaccination, and proximity to other vaccines.

Notes/statistical tests:

Can create a linear regression model that is fitted for the given independent variables and predicts the antibody response at a given timepoint (for ex., at 3 months after dose 2.). Need to make sure to codify categorical variables as dummy variables

(<https://www.moresteam.com/whitepapers/download/dummy-variables.pdf>).

Question: For which time points should I do this for?

Answer: The timepoints I choose could be informed by the ANOVA. Analysis, for example, at peak antibody response.

Question: What should the independent variables be? (by reading the objective, seems like I’m missing some things)

Answer:

1. Time between dose 1 and 2. Can be categorical (A vs. B as defined in spreadsheet, the two different timelines) or can be continuous (time between doses exactly)
2. Age
3. Gender
4. Heme malignancy? (How many different types are there; would it be realistic to codify this as a categorical variable?) There can be a lot, probably up to 10.
5. Performance status (ECOG: 0-4)
6. Treatment received in proximity to vaccination:
7. type of monoclonal antibody (Rituximab, Obinutuzumab, Daratumumab)
8. Stem cell transplant and/or CART therapy and the subsets of therapy (Y/N categorical variable, and the type of the therapy will be important)
9. (Don’t know what information to take from “Anti-cancer Therapy in the past 2 years (most recent to oldest)” Will be more than 20 therapies. Will also have data on when those therapies were done, i.e. before dose 1, between dose 1 and 2, or after dose 2.
10. Was anti-cancer treatment altered to facilitate COVID-19 vaccine (Y/N)
11. Type of vaccine (Pfizer, Moderna, P + M, AstraZeneca,
12. Diagnosed with Covid-19 (Y/N). Can create 3 categorical variables for the 3 timepoints: before vaccine, after first dose, after second dose.

For each of these independent variables, can do a univariate analysis (t tests or contingency tables), to help motivate the multivariate analysis)

Could also create a logistic regression model that fits independent variables to the outcome of being diagnosed with COVID after full treatment of vaccine (after 2 doses). The independent variables are identical as above, except now the “Diagnosed with Covid-19 (Y/N)” variable is the dependent variable (i.e., the outcome) and the Antibody concentration at a given timepoint is an independent variable.

1. To evaluate variables to vaccine response based on hematologic diagnosis and treatment exposure

Notes/statistical tests:

Question: I think I addressed this with the linear regression model?

Answer: Yes

1. To assess the durability of post-vaccination response

Notes/statistical tests:

This was effectively computed in the primary objective methodology. To assess the durability of the post vaccination response, can look at the post-hoc tests comparing the antibody concentrations post 2nd dose. If there isn’t enough statistical evidence to suggest that the means are different, then the durability is good (assuming antibody concentration even went up at all), and if there is enough statistical evidence to suggest that the means are decreasing through time, then the durability isn’t good.

1. To determine the sensitivity and specificity to a recently developed latex agglutination assay

Notes/statistical tests:

Question: Will I analyze this as well. If so, I’ll need to know what information is being collected and how is it’s accuracy being assessed (being compared to what, ELISA? PCR?)

Answer: Most likely to be completed by our co-investigators since they created this assay.

1. To determine the cellular immune responses to COVID 19 vaccine in patients treated/recently exposed to Rituximab

Notes/statistical tests:

Question: What kind of data is being collected, IFN gamma concentration? If yes, can do a T-test to compare mean concentrations of IFN gamma in rituximab patients vs. other patients. I would also suggest increasing the number of control patients to 20 as well. Would be easier to establish statistical significance.

Answer: Probably concentration. Having trouble getting control individual data.

## *Anticipated Results*

1. Valuable information on the dynamics of antibody development to the COVID-19 vaccine, including the time for peak antibody production and duration in patients with hematologic malignancies. (good)
2. Data linking antibody response to sex, age, underlying hematologic condition including therapies received for their hematologic conditions. (good)
3. The simultaneous measurement for nucleocapsid antibodies will help us assess the population penetrance of the pandemic and study the effect of previous virus exposure to the efficacy of the vaccine by comparing the strength of spike antibody response between individuals that are negative (no prior infection) or positive for anti-nucleocapsid antibodies.

Notes/statistical tests:

Population penetrance of pandemic = # of people that got covid prior to vaccine / total # of people.

My linear regression model already considers the effect of previous covid infections. However, I could also do a t-test which compares the mean antibody concentration of people with no prior infection (at a given timepoint) vs. the mean antibody concentration of people with prior infection (at the same timepoint).

On a separate note: I could repeat this same analysis and compute T-tests for all the independent variables (at a given timepoint). Slightly redundant within the context of the linear regression model, but still a possible relevant calculation.

1. Validation of a POC antibody test developed in London by us that are based on antibody-dependent agglutination of latex particles. (discussed previously)
2. Provide valuable information on the effectiveness of the spike mRNA-targeted vaccine to SARS-CoV-2 variants. This data will help guide the design of future vaccine formulations against these variants.

Notes/statistical tests:

Question: How are we assessing effectiveness against COVID variants? We’re measuring antibody and IFN gamma response, but how does that necessarily imply immunity. Further, are we trying to distinguish between effectiveness of the vaccine against the original covid vs. mutated variants?

Answer: We are assessing effectiveness by the number of people getting covid after vaccination. And yes, we are looking. Into the different variants.

6. Comparison of cellular immune responses to COVID-19 vaccine in patients with hematologic malignancy diagnosis treated with rituximab with similar diagnoses and chemotherapy without rituximab. (good)

# **Sample size determination**

**Primary objective**: proportion of our sample group (cancers) who got covid post vaccination.

Where: n is the adequate study sample to produce the desired power of the study

is the z score and alpha is the type 1 error.

is the z score and beta is the power of the study.

And ES:

Using the paper on the Ontario case report up to August 7th, p0 can be estimated as 0.00026 or 0.026%.

= 1.96 for an alpha of 0.05.

= 0.84 for a power of 80%

Assuming we have n = 412 patients in our study, we can calculate p1 and determine the minimum difference between p1 and p0 required for our result to be statistically significant.

P1 was calculated to be 0.002484 or 0.2484%

**Secondary Objective:** COVID-19 antibody concentration post 2nd vaccine.

Where the definitions are the same as previously stated, except that ES:

Where: u represents the mean antibody concentration post vaccine

(u1 is the alternative hypothesis while u0 is the null hypothesis).

I found that u0 = 1084 U/mL. I will be using the standard deviation associated with this mean as an approximation of the true standard deviation. I am currently waiting on an email back from the authors to get this value.

**Sub-study:** T-cell IFN gamma responses of target sample population (receive rux.) vs. control (receive other cancer treatments) post 2nd covid vaccine.

Where the definitions are the same as previously stated, except that ES:

Where: u1 represents the mean of the control group, and u2 represents the mean of the target

population. I currently need to find a paper that gives me the values u1 and sigma.

**Additional Questions/Concerns**

* Column K in the excel spreadsheet: “Any monoclonal antibody within 1 year of first dose?” is this referring to whether they received cancer treatment or? Because it’s right beside “Type” which says the type of chemotherapy. Answer is yes.

# **Sources:**

Repeated measures ANOVA

* Paper on the theory: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3737450/>
* Using r: <https://www.r-bloggers.com/2021/04/repeated-measures-of-anova-in-r-complete-tutorial/>
* Using r: <https://rcompanion.org/handbook/I_09.html>

Dummy variables

* <https://www.moresteam.com/whitepapers/download/dummy-variables.pdf>).

Sample size calculations:

<https://sphweb.bumc.bu.edu/otlt/mph-modules/bs/bs704_power/bs704_power_print.html>

<https://www.publichealthontario.ca/-/media/documents/ncov/epi/covid-19-epi-confirmed-cases-post-vaccination.pdf?la=en> – Proportion of healthy vaccinated people that contracted covid-19 (page 3, highlight #1)

<https://ashpublications.org/blood/article/137/23/3165/475742/Efficacy-of-the-BNT162b2-mRNA-COVID-19-vaccine-in> Mean antibody conc. In healthy vaccinated individuals at a given timepoint (look under serologic response)