Supplement

Here we describe additional details about the methodology used.

## Dataset

The Canadian Longitudinal Study on Aging (CLSA) is a comprehensive national long-term study designed to track approximately 50,000 individuals aged 45 to 85 years at recruitment for a minimum of 20 years [1]. During the COVID-19 pandemic, several studies were conducted to collect and analyze blood samples from CLSA participants, aimed at comprehending and evaluating the pandemic’s population-level health impact.

The dataset utilized in this study was sourced from the CLSA COVID-19 Antibody Study, with over 19,000 CLSA participants across ten provinces included [1]. The dataset comprises demographic information, self-reported health status, vaccination and hospitalization history, and the results of SARS-CoV-2 immunoassay tests.

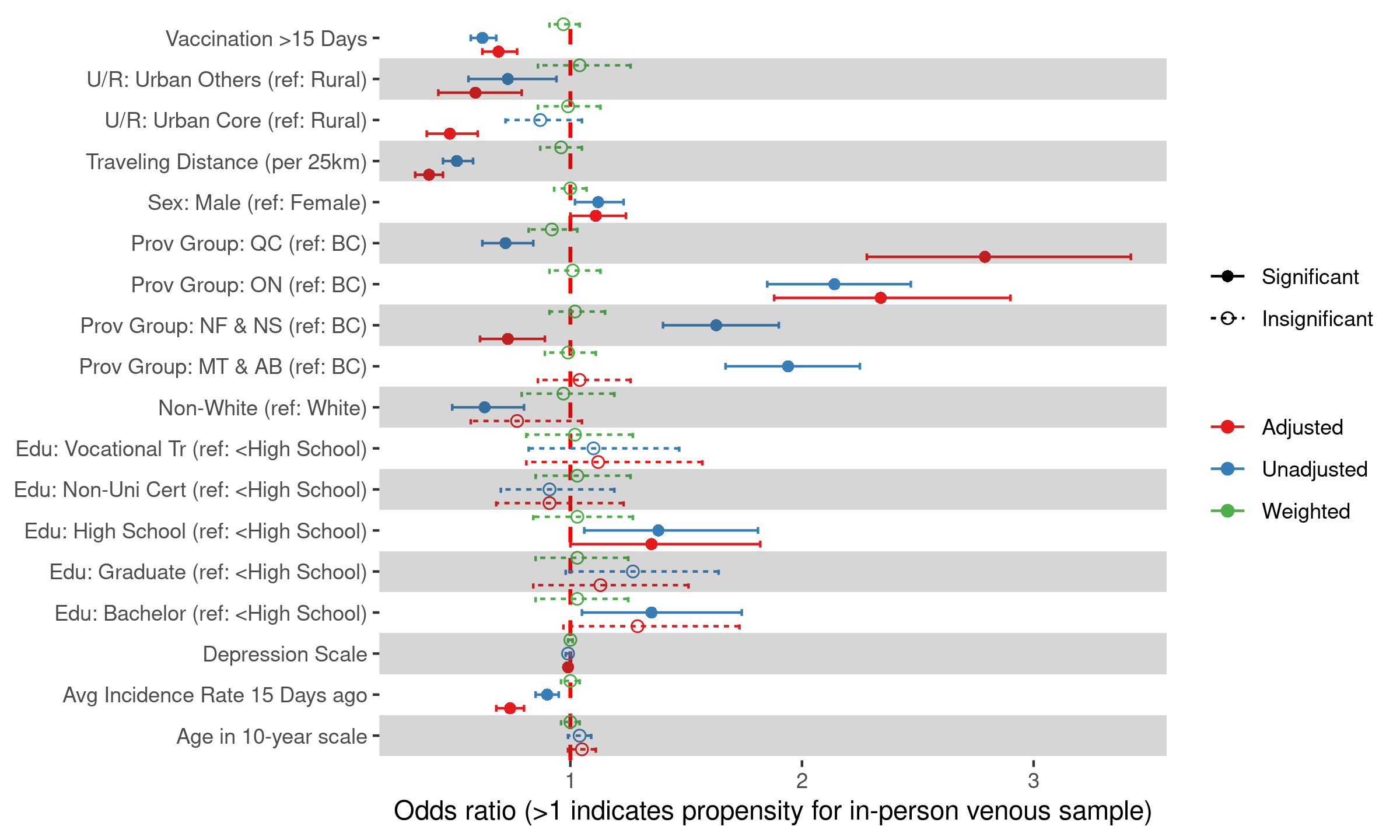
The CLSA COVID-19 Antibody Study has two complementary cohorts. The first one is called “comprehensive cohort”, in which 17,399 participants were invited and 9,753 of them agreed to participant and provided testable samples with complete results. Within this group, N=4,258 contributed venous blood samples, and N=5,495 provided dried blood samples. Participants in the comprehensive cohort were randomly selected from locations within 25-50 km of the data collection sites (DCS) and underwent in-person interviews, providing detailed information on demographics, physical and mental health status, as well as vaccination and hospitalization history. The second cohort, known as the “tracking cohort,” was interviewed by telephone. For our analysis, we will exclusively utilize data from the comprehensive cohort, as it provides more comprehensive demographic and health condition data. Additionally, we excluded participants residing in regions located more than 50 km from the data collection center, as their extended travel distance is caused by recent relocations subsequent to enrollment in the study. The final sample size included in our analysis was 7,230.

[Can we cite the CLSA flow chart here? or is it OK to make our own version of it and put it here?]

## Bivariate Analysis and Multivariate Regression Modeling

Before conducting propensity score weighting to calculate weights for each observations and analyze the weighted data, we performed an initial unadjusted bivariate analysis to compare the DBS and VBS groups. Additionally, we developed multivariate logistic regression models, regressing VBS/DBS against demographic factors, physical and mental health conditions, travel distance, vaccination status, and time since the last COVID outbreak. The comprehensive results generated by those simple logistic regression models and the full model are displayed in the odds ratio plot (**Fig. 1**).

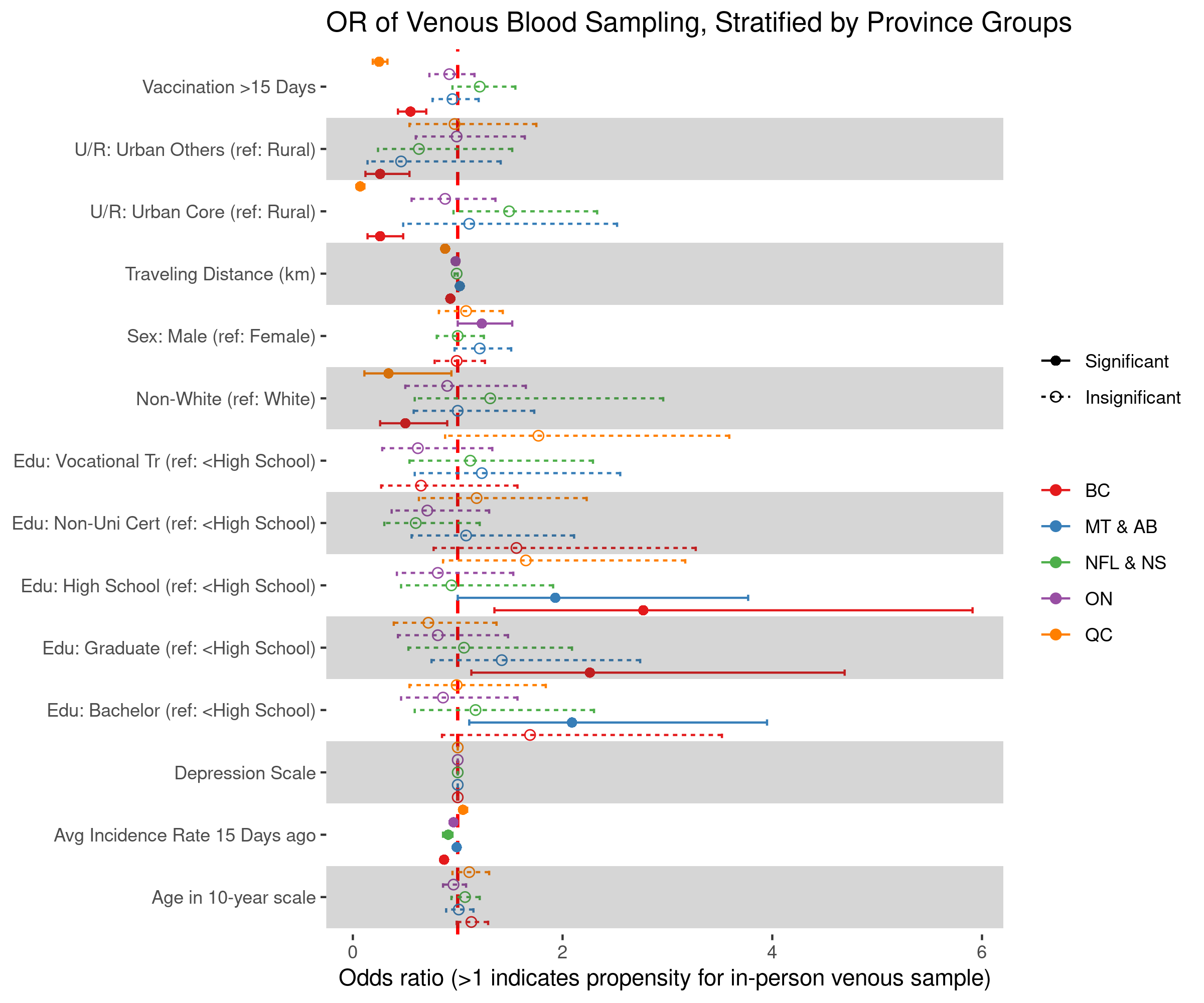
**Figure 1**: Odds ratios from the propensity score model predicting who is more likely to give an in-person venous sample vs. mailed DBS sample



Upon completion of the full model, we constructed a best-fit model utilizing a backward stepwise model selection strategy. The model with lowest Akaike Information Criterion (AIC) value was selected as the best-fit model. Subsequently, we applied this best-fit model to calculate propensity scores for the selection of VBS over DBS. Participants from both groups will then be weighteded based on their propensity scores using Inverse probability of treatment weighting (IPTW). As depicted in **Figure 1**, balance has been achieved between the DBS and VBS groups. All factors used in the propensity score model distribute evenly across the two comparison groups.

Based on the national data, we conducted a stratified analysis by province. Our study sample included participants from seven provinces, which were recategorized into five groups according to their locations and geographic proximities: British Columbia (BC), Manitoba (MB) & Alberta (AB), Ontario (ON), Quebec (QC), and Newfoundland (NL) & Nova Scotia (NS). The grouping is based on their geographic locations and proximity to each other. The pandemic’s impact varied across these regions, with residents adapting differently to the diverse policies implemented by their respective health departments [add citations]. As illustrated in **Figure 3** below, the odds ratios for venous blood sampling vary among the five provincial groups.

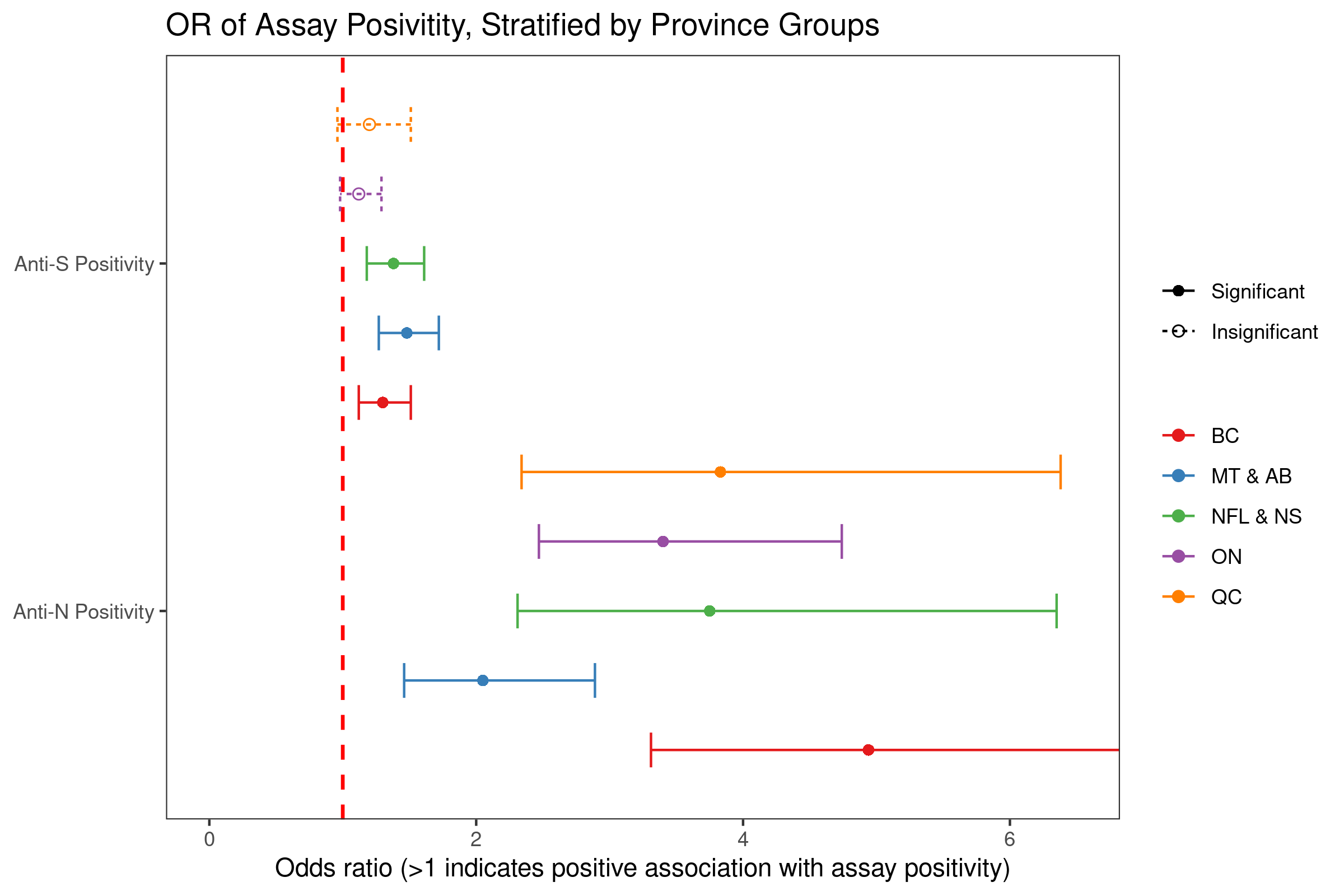
**Figure 3**: Odds Ratio of Venous Blood Sampling, Stratified by Province Groups



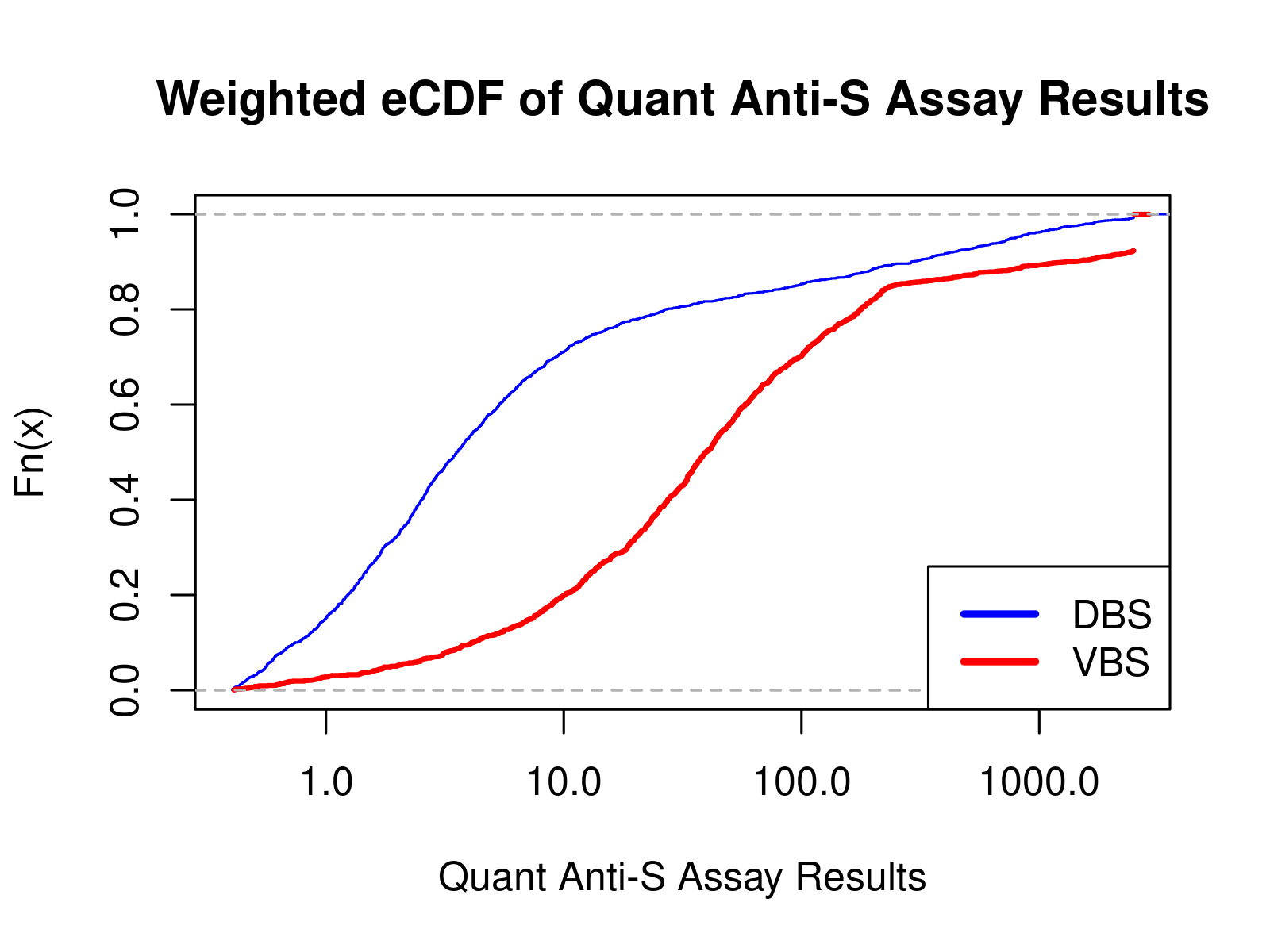
## Analysis with Weighted Data

Utilizing weighted data, we further analyzed the differences in assay positivity between the DBS and VBS cohorts. The outcomes of the stratified analysis align with those from the unstratified national dataset. We observed an increased odds of assay positivity in venous blood samples for both anti-N and anti-S assays.

**Figure 4**. Odds Ratio of Assay Positivity, Stratified by Province Groups



To analyze the variation in quantitative Anti-S immunoassay results, we employed the weighted empirical cumulative distribution function (eCDF) to visualize the distribution in the DBS and VBS cohorts (**Fig. Y**). We used Kolmogorov-Smirnov (KS) Test to assess whether the distributions are statistically-significantly different.



# References

1. Raina P, Wolfson C, Kirkland S, et al. Cohort Profile: The Canadian Longitudinal Study on Aging (CLSA). *International Journal of Epidemiology*. 2019;48(6):1752-1753j. doi:[10.1093/ije/dyz173](https://doi.org/10.1093/ije/dyz173)