**GWAS Analysis Plan for Hypoglycemia**

**Objective:** To conduct a genome-wide association study (GWAS) of hypoglycemia in those who have been diagnosed with diabetes in the Million Vet Program (MVP) study population. The study will determine the odds of having hypoglycemia given the genotypes at specific loci.

**Data:** In our GWAS, participants must meet all these criteria to be included in the study:

* They had one primary care diagnosis of diabetes or 2 or more diagnoses anywhere
* They filled at least one medication prescribed for diabetes
* They had at least one random plasma glucose (RPG) measurement after diabetes diagnosis
* Their diabetes diagnosis (FIRST\_DIABETES\_ DT) was before enrollment in MVP (ENROLLED\_DATE)

**Methods:** We will conduct the GWAS using a case-control study design. A case is defined as a participant that met the inclusion criteria and has hypoglycemia: received a diagnosis from an Emergency Department (ED) visit (ED\_HYPO\_DX\_FLAG) or an outpatient RPG measure < 70 (GLUCOSE\_LT70\_FLAG) at any point after diabetes diagnosis. A control is defined as a participant that met the inclusion criteria, did not meet the case definition, and had a Primary Care Physician (PCP) visit in every calendar year since diabetes diagnosis or had a glucose reading in every calendar year since diabetes diagnosis or had >= 1 Insulin or Sulfonyurea fill any time after diabetes diagnosis. After the assessment of the inclusion criteria, our study population consists of 114,096 participants: 38,114 cases and 75,982 controls.

The GWAS study will be performed using a logistic regression model to associate case status (CASE\_CONTROL\_FLAG) with genotype data gathered using the Affymetrix MVP Array for each participant. We will use an additive model of the allele count: 0, 1, or 2, at each locus as the main predictor in our model. We will adjust the model for age at enrollment (AGE) in MVP, sex (GENDER), and the top 10 principal components (PCs) of ancestry. The analysis will be conducted using a simple stratification of Black and White (self-report or can we use HARE?) to account for African and European ancestry. PCs will be calculated separately for these two strata.

We will look for significant variants using a Bonferroni-corrected threshold of 5x10-8.

Questions (leaving this in as we need input from Yan): Do we need to calculate the PCs ahead of time or will that be part of the GWAS? Do we need to do any SNP trimming (Hardy-Weinberg, missing SNPs, homozygous SNPs? Outliers of principal components?)

**Secondary Analysis:** As there was concern that a stricter or looser definition of Control could impact the results, we will conduct a sensitivity analysis using varying definitions of the Control participant (Table 1). These variations will be conducted:

1. Rows 2-8 in Table 1: the definition described above
2. Rows 2, 4, 6, 8: to impose a stricter definition of the Control participant
3. Row 8: the strictest possible definition that the Control participant meets all 3 criteria

Table 1: Variations of the Control Definition

|  |  |  |  |
| --- | --- | --- | --- |
| **>=1 Insulin or Sulfonyurea fill any time after Diabetes Dx** | **PCP Visit in every calendar year since Diabetes Dx** | **Glucose reading in every calendar year since Diabetes Dx** | **# of Patients** |
| 0 | 0 | 0 | 9,376 |
| 0 | 0 | 1 | 2,009 |
| 0 | 1 | 0 | 4,752 |
| 0 | 1 | 1 | 21,235 |
| 1 | 0 | 0 | 16,646 |
| 1 | 0 | 1 | 2,893 |
| 1 | 1 | 0 | 8,275 |
| 1 | 1 | 1 | 29,138 |
|  |  | **Overall Total:** | **94,324** |