**APPENDIX 1: Internal Manuscript Concept Sheet**

|  |
| --- |
| **eMERGE Network: Manuscript Concept Sheet** |
| **Reference Number** *(to be assigned by CC)* | NT459 |
| **Submission Date** | August 24, 2022 |
| **Project Title** | A polygenic risk score for white blood cell variation |
| **Tentative Lead Investigator** *(first author)* | Jonathan Mosley |
| **Tentative Lead Investigator Email *(first author)*** | Jonathan.d.mosley@vumc.org |
| **Tentative Senior Author** *(last author)* | Vivian Kawai (jonathan.d.mosley@vumc.org) |
| **All Other Authors**  | Sara Van Driest; Dan M. Roden, MD; Scott Borinstein |
| **Sites Participating** | Vanderbilt University Medical Center |
| **Background / Significance** | Variation in white blood cell counts in populations are influenced by polygenic variation. Large GWAS of WBC variation have identified numerous SNPs associated with variability in WBC cell types. |
| **Outline of Project** | This project will validate an observed association in BioVU between a white blood cell polygenic score derived from summary statistics from a large WBC GWAS (Cell. 2020 Sep 3;182(5):1198-1213.e14) and phewas code “288.1: Decreased white blood cell count” among adult subjects. |
| **Desired Data - Common Variables\*** *(Available from the CC)* | Demographics ☐ICD9/10 codes☐CPT codesCheck mark, Wingdings font, character code 252 decimal.Phecodes☐BMI | ☐Common Variable Labs☐Common Variable Meds☐ Geocoding 2015 ACS variables☐Other: Case/Control status  |
| **Other Desired Data *(Available from participating sites)*** | *Please specifically list out any data elements that participating sites would collect or extract from clinical or other sources for this project (i.e. not common variables above)*  |
| **Desired Genetic Data** | Check mark, Wingdings font, character code 252 decimal.eMERGE I-III Merged set (HRC imputed, GWAS)☐eMERGE PGx/PGRNseq data set ☐eMERGEseq data set (Phase III)☐eMERGE Whole Genome sequencing data set☐eMERGE Exome chip data set☐eMERGE Whole Exome sequencing data set☐Other (not listed above): PheWAS data |
| **Does project pertain to an existing eMERGE Phenotype?** | Check mark, Wingdings font, character code 252 decimal.Yes, if so please list: PheWAS code 288.1 ☐No |
| **Planned Statistical Analyses** | A logistic regression model, adjusting for birth decade, sex and principal components will be used to test for the association between the pheWAS code and the WBC PGS. A p<0.05 will be considered significant. |
| **Ethical Considerations** | All data will be de-identified and aggregated, thus posing no ethical concerns.  |
| **Target Journal** | Hematology |
| **Milestones***(This section should include the key dates for completion of project, including approval, project duration, draft completion, and submission.)* | September 10, 2022 – ApprovalProject duration – 1 months or sooner Draft completion – by 2 months (October 31, 2022) or soonerSubmission – by November 1, 2022 or sooner |

***\*\**** *This section should include the timeline for completion of project, including: approval, project duration, first and second draft of the paper and submission.*