Supplemental Methods

The Vitamin Intervention for Stroke Prevention (VISP) participants

The VISP trial was a multi-center, randomized, double-blind, controlled clinical trial comparing daily supplementation of high dose versus low dose folic acid, vitamin B_6 and vitamin B_{12} for reduction in recurrent cerebral infarction and nonfatal myocardial infarction (MI) or mortality.[1] The VISP trial was approved by the relevant institutional review boards (IRBs), including the IRBs of Wake Forest University School of Medicine (coordinating center) and the University of North Carolina at Chapel Hill School of Medicine (statistical center) as well as individual recruiting sites, and conducted according to the Declaration of Helsinki. All participants provided written, informed consent. VISP data analysis by the Genomics and Randomized Trial Network (GARNET) was approved by University of Virginia School of Medicine IRB.[2]

VISP enrolled individuals 35 years or older with a non-disabling cerebral infarction (NDCI) within 120 days of presentation who had a homocysteine (HCY) level above the 25th percentile.[1, 3] NDCI was defined as a clinical ischemic stroke with deficits lasting at least 24 hours or transient clinical syndrome with an infarction in the part of the brain corresponding to the symptoms demonstrated by CT or MRI imaging and no residual disability. Additionally, those with embolic stroke from a cardiac source or stroke due to operable carotid disease were excluded. The high-dose formulation contained 25 mg pyridoxine (B_6), 0.4 mg cobalamin (B_{12}), and 2.5 mg folic acid, and the low-dose formulation contained 200 µg pyridoxine, 6 µg cobalamin and 20 µg folic acid. Enrollment in VISP began in August 1997, and was completed in December 2001, with 3,680 participants (High dose=1,827 and low-dose=1,853) enrolled from 55 clinic sites across the U.S. and Canada and one site in Scotland.

The genetic sub-study enrolled and consented 2,100 individuals for their data to be used for future analyses including the current study. The VISP genetics cohort in general reflected the study population as a whole, although the sample for the genetic studies included a higher proportion of men than women (Table 2). As expected for an ischemic stroke population, the study sample has a high prevalence of diabetes and hypertension (Table 2). According to the VISP protocol, all individuals had circulating HCY levels in the top quartile (Table 2). The VISP sample consisted of 1,725 (82.1%) individuals of European

descent, 258 (12.2%) individuals of African descent and 117 (5.6%) individuals with other or unspecified race/ethnicities, based on self-reported data.

Genotyping was performed on the Illumina HumanOmni1-Quad-v1 array (Illumina, Inc.) at the Center for Inherited Disease Research, Johns Hopkins University. Quality control measures included filtering SNPs based on missing call rate, Mendelian errors in control trios, deviation from Hardy-Weinberg equilibrium in controls, discordant calls in duplicate samples, sex differences in allele frequency or heterozygosity, and minor allele frequency.[4] Data were included in the analyses for those with an overall missing call rate <2%, although sample-chromosome combinations were also excluded where a gross chromosomal anomaly was detected or when the chromosome-specific missing call rate was > 5%. Genetic imputation was used to increase the number of SNPs in the analysis, by use of the multi-ethnic phase 1 interim release from the 1000 Genomes (1000G) Project.[5] Imputation target variants were defined as those with MAF ≥ 0.005 across 629 1000G samples. Imputation was performed using BEAGLE imputation software[6] (v3.3.1), for chromosomes 1-22 and the X chromosome. The imputed dataset contained total 7,500,450 variants; 766,577 of which (10.2%) were observed from the direct genotyping. In addition to the primary imputation analysis, additional imputations were run on chromosome 22 and the X chromosome, masking a random 10% of observed SNPs to empirically assess imputation quality. The squared correlation between observed and imputed allelic dosages (dosage r^2) was used to summarize the imputation quality. The median dosage r^2 was 0.933 for chromosome 22 masked SNPs and 0.930 for X chromosome masked SNPs. The genotyped and imputed datasets, along with a detailed report on imputation methodology, is available through dbGaP (study accession number phs000343).

Stroke recurrence definition

The definition of recurrent ischemic stroke was an acute neurological ischemic event of at least 24h duration with focal signs and symptoms, without evidence of primary intracranial hemorrhage or other alternative explanation, together with one of the following: a one-point increase in the NIHSS in a previously normal section or, lacking this, an appropriate new or extended abnormality seen on CT or MRI. Diagnoses were reviewed by the local neurologist, two endpoint reviewers, and occasionally a full Stroke Endpoint Review Committee. Underlying cause of death was decided by a Death Review

Committee composed of physicians not affiliated with VISP. Decisions were based on information available from hospital records, death certificates, coroners' reports, or physicians' questionnaires. Stroke lesion and stroke subtype information are not available in VISP

Supplemental References

1. Spence JD, Howard VJ, Chambless LE, et al. Vitamin Intervention for Stroke Prevention (VISP) trial: rationale and design. Neuroepidemiology 2001;20:16-25.

2. Keene KL, Chen WM, Chen F, et al. Genetic Associations with Plasma B12, B6, and Folate Levels in an Ischemic Stroke Population from the Vitamin Intervention for Stroke Prevention (VISP) Trial. Front Public Health 2014;2:112.

3. Toole JF. Vitamin intervention for stroke prevention. J Neurol Sci 2002;203-204:121-124.

4. Laurie CC, Doheny KF, Mirel DB, et al. Quality control and quality assurance in genotypic data for genome-wide association studies. Genet Epidemiol 2010;34:591-602.

5. Altshuler D, Gibbs R, Peltonen L, et al. Integrating common and rare genetic variation in diverse human populations. Nature 2010;467:52-58.

6. Browning B, Browning S. A unified approach to genotype imputation and haplotype-phase inference for large data sets of trios and unrelated individuals. Am J Hum Genet 2009;84:210-223.