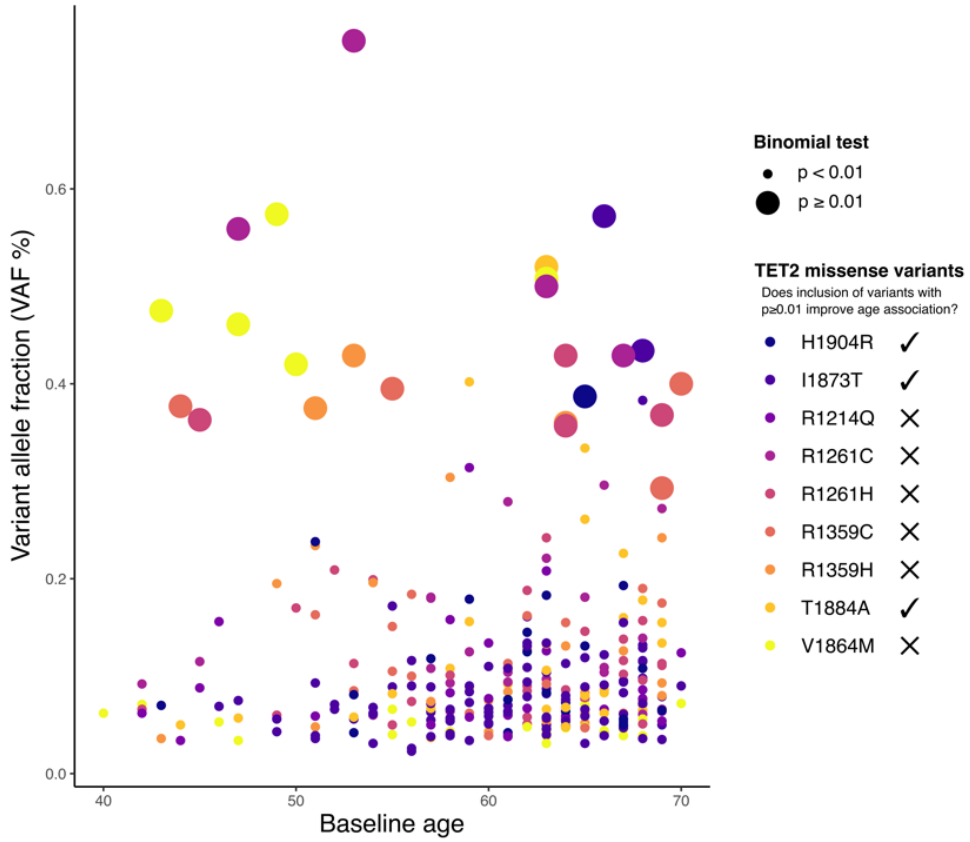
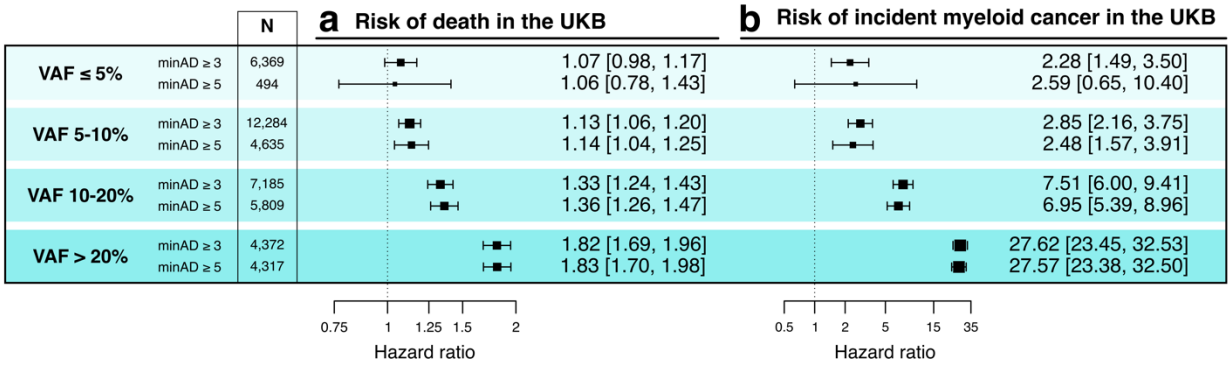
Supplemental Material

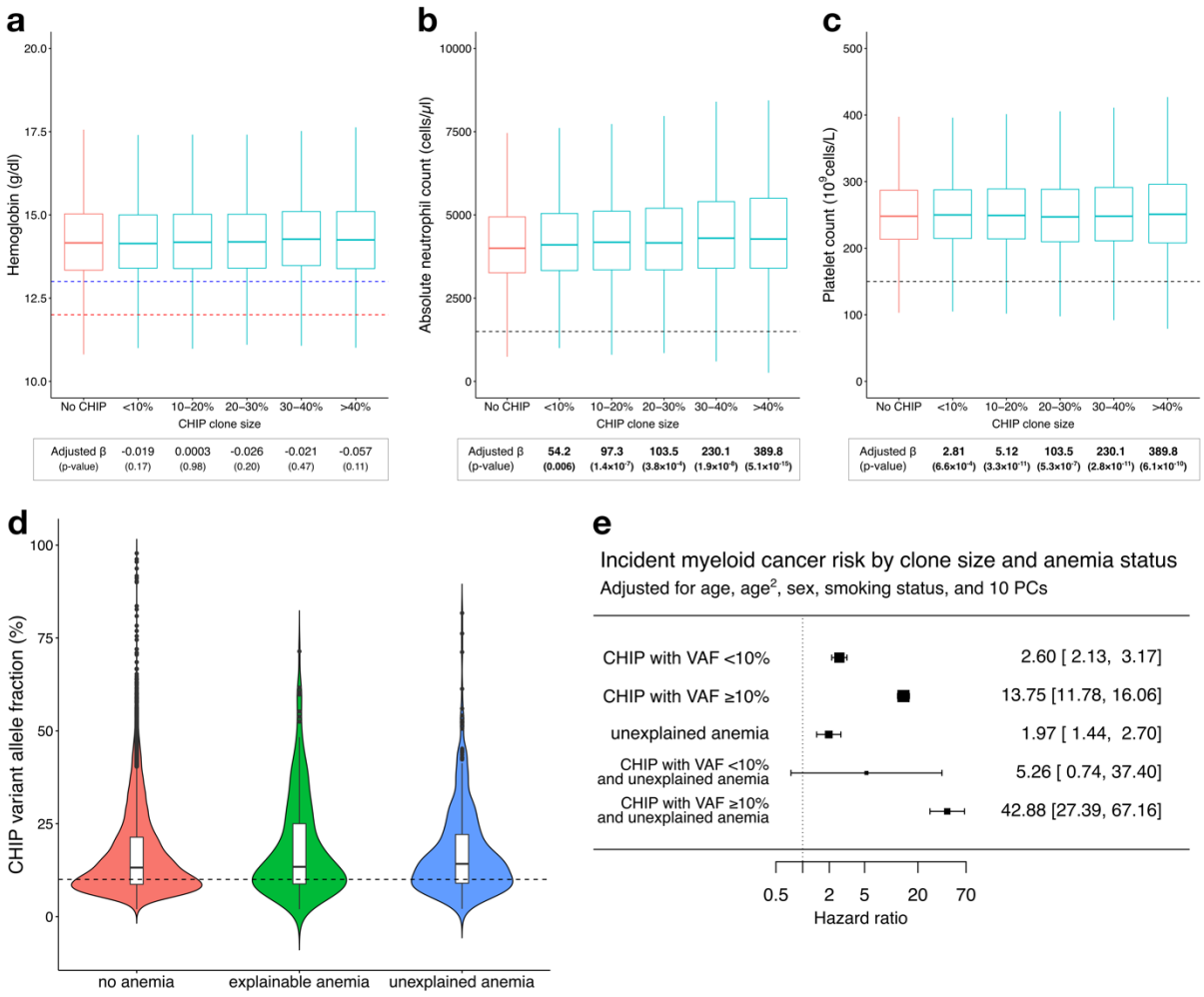
1. **Supplemental Figure 1** – Distribution of variant allele fractions (VAF) by age for *TET2* missense variant groups where only a portion of variants fail the binomial test.
2. **Supplemental Figure 2** – Association with death and incident myeloid cancers for CHIP defined by a minAD threshold of 3 compared to a minAD threshold of 5.
3. **Supplemental Figure 3** – Association of CHIP with hematologic traits stratified by clone size and identification of possible CCUS cases.
4. **Supplemental Methods**



Supplemental Figure 1. Distribution of variant allele fractions (VAF) by age for *TET2* missense variant groups where only a portion of variants fail the binomial test. Cases where addition of variants failing the binomial test improves the association of VAF with age are indicated.



Supplemental Figure 2. Association with death and incident myeloid cancers for CHIP defined by a minAD threshold of 3 compared to a minAD threshold of 5.



Supplemental Figure 3 – Association of CHIP with hematologic traits stratified by clone size and identification of possible CCUS cases in the UKB. CHIP is not associated with differences in baseline hemoglobin levels (**a**) but is associated with higher absolute neutrophil and platelet counts (**b, c**) across variant allele frequency (VAF) strata, including in regression models adjusted for age, age^2, sex, smoking status, and 10 principal components of genetic ancestry. Dashed lines indicate the typical lower levels of normal for these hematologic parameters. Among individuals with CHIP, the VAF distribution among those with unexplained anemia was not different compared to those with potentially explainable anemia or no anemia, and the proportion of small clones (VAF <10%) was not statistically different (**d**). Unexplained anemia increased the risk of incident myeloid cancer by approximately two-fold among those with small clones and three-fold among those with large clones (**e**).

Supplemental Methods

Simulation testing

We performed simulation testing in order to estimate the amount of variant misclassification present when using different minAD strata. We first identified that a minAD ≥ 5 had the greatest association with age and the *TERT* promoter variant in the UK Biobank. In our simulations, we replaced subsets of the CHIP call set defined by minAD ≥ 5 with randomly selected CHIP-free individuals from the cohort. For example, we replaced 5% of the dataset with CHIP-free individuals to estimate the effect that 5% of misclassification would have on the age- and *TERT* promoter variant associations. We carried out these simulations with 5%, 10%, 20%, 25%, 30%, 40% and 50% sample replacement, and each simulation was performed 20 times.

Hematologic phenotype definitions

Hematologic parameters from bloodwork at enrollment were used to define anemia (hemoglobin <12g/dL for males and <13g/dL for females), thrombocytopenia (platelet count of $<150 \times 10^6$ cells/L) and leukopenia (1.8×10^6 cells/L). “Possibly explainable anemia” was defined as an individual with either an ICD10 code for an anemia of known etiology (i.e., D50-D61, D63, D64.0-8) or a record of iron or vitamin B12 supplementation on enrollment. The ICD10 codes D62 (acute post-hemorrhagic anemia) and D64.9 (anemia, unspecified) were not included in this definition.