Liquid biopsy is a powerful, non-invasive tool for profiling tumors and identifying clinically relevant variants. The presence of clonal hematopoiesis (CH) variants, and biological noise, due to aging and therapy has potential to confound biomarker interpretation.

Currently, comprehensive methods to filter non-tumor variants require genotyping the white blood cell (WBC) fraction of the paired plasma sample, which is a costly, complicated workflow.

A plasma-only, bioinformatics solution to identify non-tumor variants is needed for accurate biomarker assessments in the cell-free DNA (cfDNA).

**Methods**

Variant calls were obtained from >250,000 plasma samples comprising healthy donor, early and late-stage cancer patients sequenced on the Guardant360™, GuardantREVEAL™, GuardantOMNI™ and GuardantInfinity™ liquid biopsy panels as well as public tissue datasets.

The model was trained on paired plasma and WBC datasets and optimized with 10-fold cross-validation to produce a non-tumor and tumor variant classifier. To validate these calls, an independent cohort of 72 paired plasma and WBC advanced cancer samples were genotyped on the GuardantInfinity™ assay. A cohort of 76 healthy donor samples, genotyped on the GuardantOMNI™ assay was also assessed.

**Results**

**Model performance:** High ROC AUC and accuracy for predicted calls

A. Validation: Late-stage cohort of 72 paired plasma and WBC samples

B. Validation: Healthy donor cohort of 76 paired plasma and WBC samples

**Feature importance:** Assay-specific engineered features among most important

A. Top 10 features

B. VAF Features

C. Cancer uniformity feature

**Conclusions**

Our bioinformatic model exhibits high sensitivity and specificity with WBC for discriminating tumor and non-tumor using only cfDNA.

Our bioinformatic model has improved sensitivity for identifying non-tumor variants over WBC sequencing at low VAFs (>0.6%).

In a paired plasma and WBC late stage cancer cohort, the majority of non-tumor variants were in known clonal hematopoiesis genes and variants of uncertain significance. No clinically actionable variants, except in ATM and CHEK2, were confirmed or annotated as non-tumor.