Cell-Free DNA methylation as a biomarker of malignant vs. benign lung nodules: Cross-validation of 2022 and 2023 patient cohorts

Sin Danju (SEAS 2026)\textsuperscript{1}; Prishna Martinez (COL 2025)\textsuperscript{1}; Jacob E. Till, MD, PhD\textsuperscript{2}; Zacharya Yazdani\textsuperscript{3}; Melinda Yin, PhD\textsuperscript{4}; Jeffrey Thompson, MD\textsuperscript{5}; Erica L. Carpenter, MBA, PhD\textsuperscript{6}\\
\textsuperscript{1}Division of Hematology, Department of Medicine, Perelman School of Medicine, University of Pennsylvania; \textsuperscript{2}Division of Pulmonary, Allergy, and Critical Care Medicine, Department of Medicine, Perelman School of Medicine, University of Pennsylvania.

**Background**

- A lung nodule is a rounded lesion which is measured to be up to 3 cm in size.\textsuperscript{1}
- Screening with Low-dose CT (LDCT) scans has shown to be effective in detecting potentially malignant lung nodules (≤ 3 cm diameter) and decreasing lung cancer mortality.\textsuperscript{1}
- Small nodule classification based on morphologic alone can be difficult, leading to an "indeterminate" diagnosis.\textsuperscript{1}
- LDCT scans can lend false-positive results for indeterminate nodules and repeat screening may be required, increasing radiation exposure risk.\textsuperscript{1}
- Tissue biopsy is the current standard for molecular detection, however bronchoscopy is an invasive procedure and re-biopsies are not always possible.

**Methods**

- cfDNA and ctDNA
  - Cell-free DNA (cfDNA) can be found in the blood of both healthy individuals and patients with cancer, and is often found in higher concentrations in patients with cancer.\textsuperscript{2}
  - A portion of cfDNA derived from circulating tumor cells in patients with cancer is known as circulating tumor DNA (ctDNA).\textsuperscript{2}
  - ctDNA can serve as a non-invasive clinical tool to detect cancer.\textsuperscript{2}

- cfDNA Methylation Analysis in Patients with Lung Nodules:
  - CpG Methylation is the addition of a methyl group to the cytosine in a cytosine-guanine nucleotide sequence. This methylation tendency to "silence" expression of a gene.
  - Methylation signatures may help detect malignant versus benign lung tumors. Unmethylated CpG sites can be identified after treatment with the APOBEC enzyme which converts CpG to Cpg, allowing us to determine which regions are methylated, a change detected by sequencing.
  - Hypermethylated sites at tumor suppressor gene promoters could indicate malignancy because they “silence” the mechanisms that are intended to suppress rapid growth of tumorous cells.
  - ctDNA is a combination of extra-cellular DNA from various cell types, which can detect cancer contribution by referencing unique methylation signatures of individual cell types as corresponding to the amount of DNA in our sample: this is deconvolution.

**Results**

- D. Supervised Clustering of Cohort 2022

**Discussion and Future Directions**

- The supervised clustering of the DMRs for both the 2022 Cohort and 2023 Cohort worked well to separate malignant vs. benign patients. However, when cross-validating the 2022 Cohort data with the DMRs found for the 2023 Cohort, there is no clear separation established between malignant and benign patients. Similarly, when the 2022 Cohort was clustered using the DMRs found in the 2023 Cohort, there is also no clear separation between malignant and benign patients. This suggests that our 2022 and 2023 models might be overfit and not generalizable to new data.
- There were only 2 similar CpG sites between 2022 Cohort and the 2023 Cohort for differentially methylated regions. The EPIC v2.0 Array was used for the methylation profiling of the 2023 Cohort, while the EPIC v1.0 Array was used for the methylation profiling of the 2022 Cohort.

**Acknowledgements**

We would like to thank the entire Carpenter Lab for helping and supporting us through our research, especially Zach for extensively guiding us through the methylation data analysis and Dr. Till and Dr. Yin for helping us understand this project. We also wish to express our sincere gratitude to Dr. Carpenter and Dr. Thompson for the opportunity to work in their lab. Funding was provided from CURE through the Penn Undergraduate Research Mentoring Program.

**References**


**Author:** reducing artifactual detection of DNA methylation by Infinium cfMeDIP sequencing of circulating tumor DNA (ctDNA) can be used as a biomarker for monitoring disease progression and response to therapy. J Cell Mol Med. 2018;22(9):4543-4553. doi:10.1111/jcmm.14293


