

# FLAMINGO: Accurate cancer detection from ultra-low-pass whole genome sequencing of cell-free DNA

D.C.L. Vessies, T. Bucho, E. Post, K. Roohollahi, J. van den Berg, A. Foster, S. Patel, A. Vandor, D.J.S. Makarawung, F. van Asch  
RenovaroCube, Amsterdam, The Netherlands



## Introduction

- Cell-free circulating tumour DNA (ctDNA) is a minimally invasive biomarker
- Current cancer detection methods are insufficient for early cancer detection
- RenovaroCube develops an AI Platform for cancer diagnostics
- No single model achieves the requisite sensitivity when tuned for high specificity
- The AI Platform bridges this gap with:
  - Multi-omic data ingestion
    - Methylation, CNV, gene expression, mutation, fragmentomics
  - Data type (array, sequencing platform) -agnostic approach
  - Comprehensive biomarker mining
  - Library of trained models

## Flamingo

- One model from the library of trained models in the AI Platform
- Fragmentomics-based:
  - cfDNA fragment lengths
  - cfDNA breakpoint sequence motifs

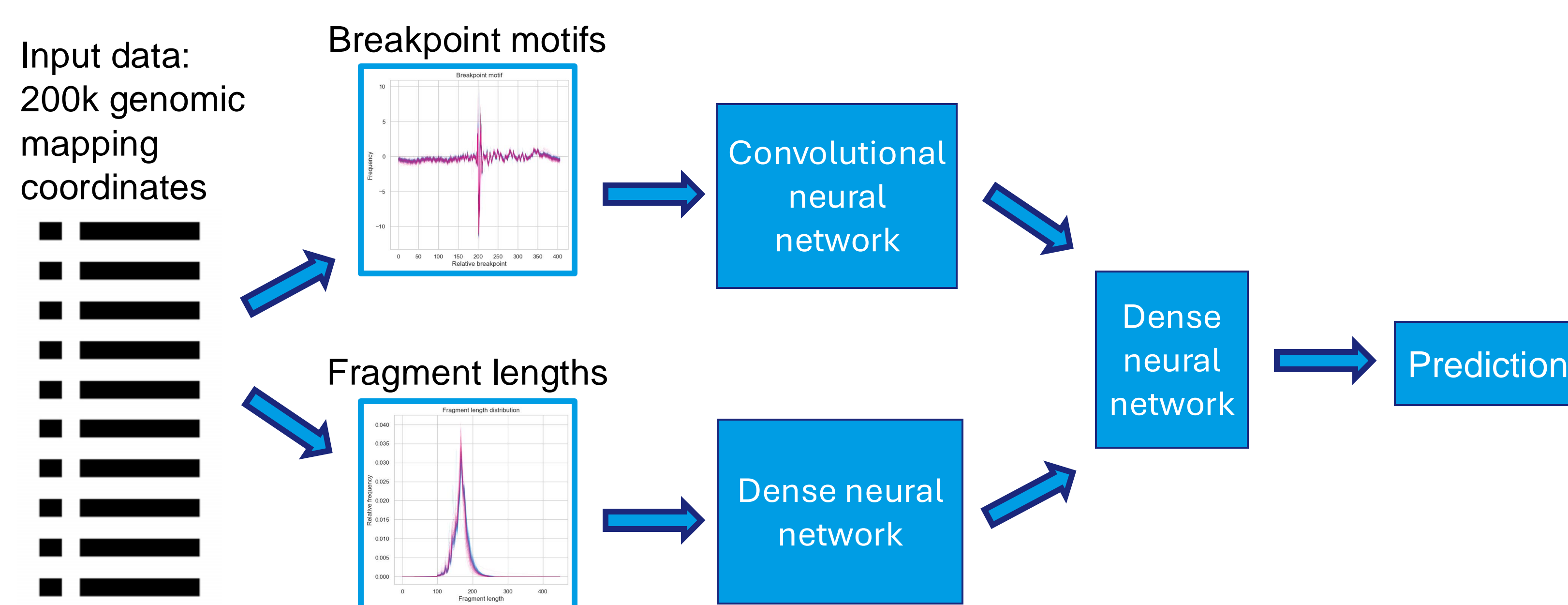


Figure 1. Schematic representation of Flamingo neural network architecture.

## Methods

- Publicly available Delfi data <sup>1,2</sup>
  - Only 200,000 reads/fragments per patient (~0.01x coverage)
- Training set (data not shown):
  - Healthy donors (n=136)
  - Cancer patients (n=127) [7 types of cancer]
- Test set:
  - Healthy donors (n=48)
  - Cancer patients (n=33) [same 7 types of cancer]

## Results

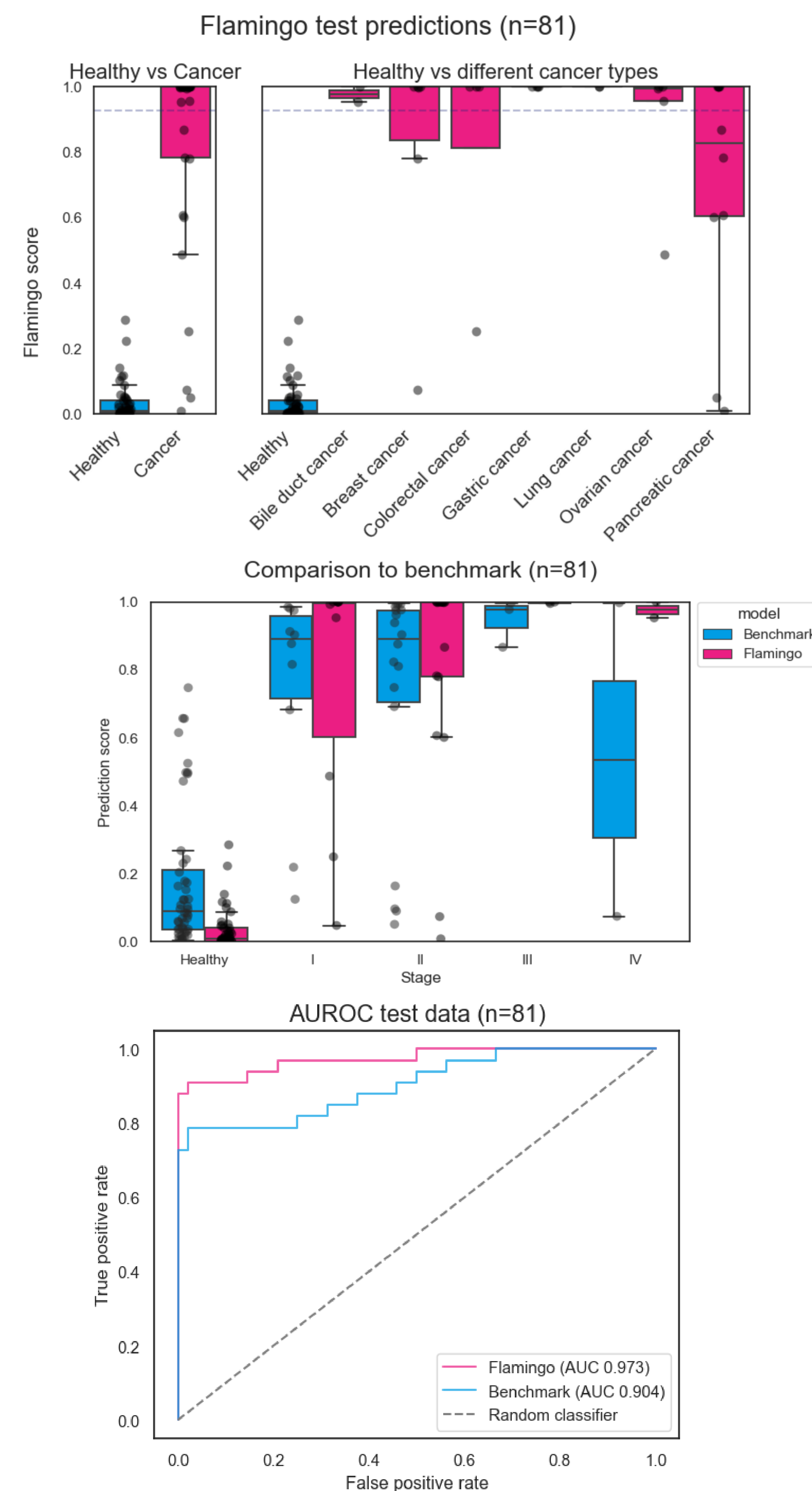


Figure 2. Flamingo predictions on unseen test set. **Top.** Across different cancer types. **Middle.** Comparison to benchmark, across cancer stages. **Bottom.** Receiver-operator curve. AUC = Area Under Curve.

## Flamingo vs benchmark, non-overlapping predictions

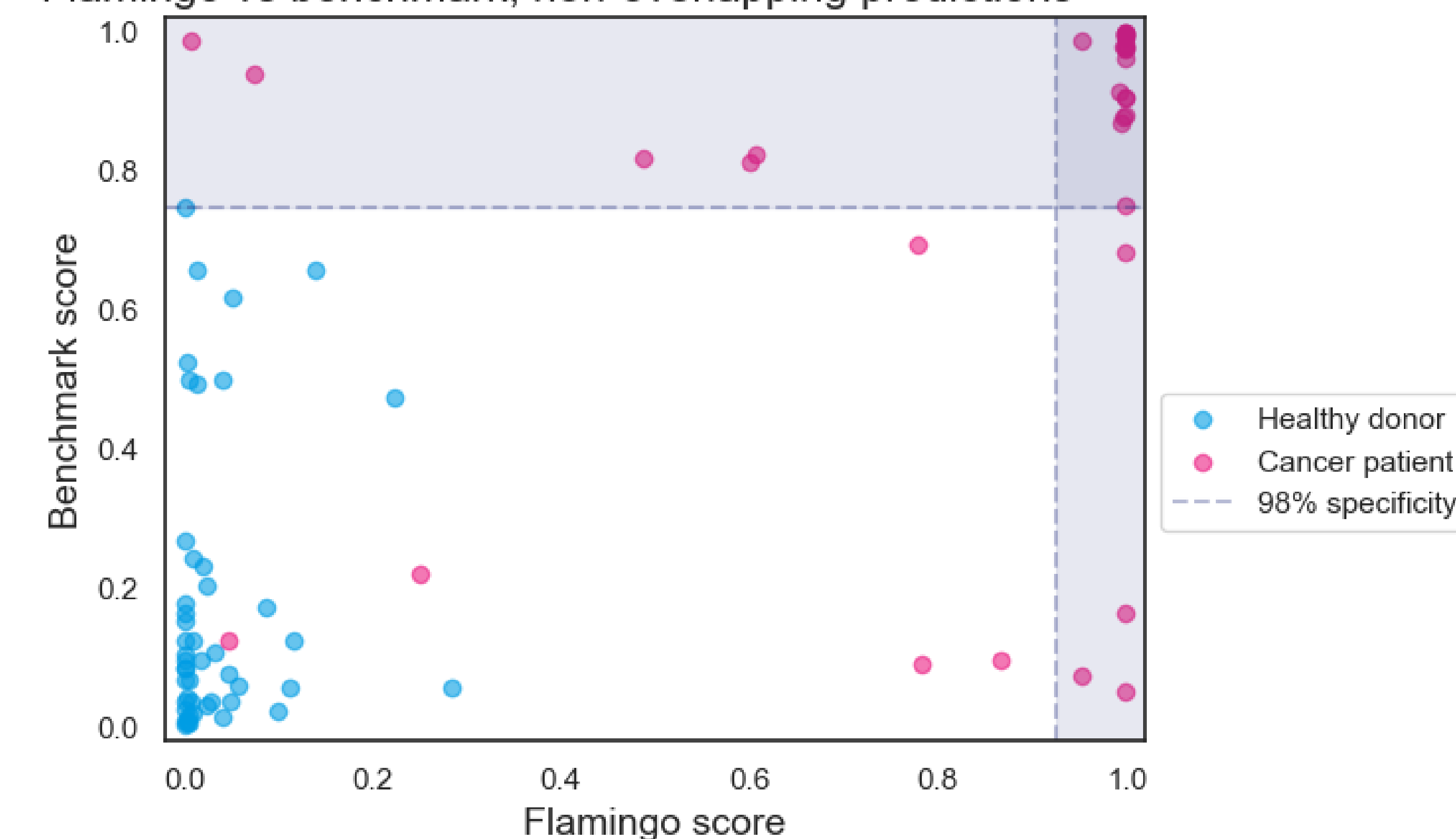


Figure 3. Flamingo vs benchmark results. Shaded areas indicate results greater than the 98% specificity cut-off as established in the training set.

## Non-overlapping cases detected

- Flamingo and benchmark each identified cases that the other missed:
  - Flamingo positive (bottom-right) [n=4]
  - Benchmark positive (upper-left) [n=5]
- No correlation with cancer type or cancer stage
- Joint sensitivity: 85.3%
- Diverse types of cancer signal from same input data indicate:
  - No single model can detect all cases
  - Multi-model approach may help to detect more cases

## Conclusions

- We develop a multi-omics AI Platform for cancer diagnostics
- Flamingo is a fragmentomics-based model, part of the AI Platform
- Flamingo showed similar performance to benchmark method (Delfi) while requiring less data
- Non-overlapping cases were identified
  - This suggests a multi-model approach

## References:

1. Cristiano S, Leal A, Phallen J, Fiksel J, Adleff V, Bruhm DC, et al. Genomewide cell-free DNA fragmentation in patients with cancer. *Nature*. 2019 Jun 29;570(7761):385-9.
2. Zheng H, Zhu MS, Liu Y. FinaleDB: a browser and database of cell-free DNA fragmentation patterns. *Bioinformatics*. 2021 Aug 25;37(16):2502-3.

## Corresponding authors:

[jasper@renovarocube.com](mailto:jasper@renovarocube.com), [daan@renovarocube.com](mailto:daan@renovarocube.com)

Link to Flamingo whitepaper:

