The Circulating Cell-free Genome Atlas (CCGA; NCT02889978) study is a prospective, multicenter, observational, case-control study with longitudinal follow-up to support the development of a plasma circulating cell-free DNA (cfDNA)-based multi-cancer early detection test.

We previously reported that a targeted methylation (TM) assay and machine-learning classifier were able to predict cancer with a single fixed false positive rate of <1% and simultaneously determine tissue of origin with >90% accuracy.1-3

Previously we showed that the fraction of cfDNA fragments from tumor versus normal cells (tumor fraction, TF) is a stronger predictor of cancer detection than clinical stage and was an equivalent predictor of survival.4

Here, we used colorectal cancer (CRC) as one example to understand the biophysical determinants of TF in solid tumors.

The model included tumor surface area (TSA) as one example to understand the biophysical determinants of TF in solid tumors.

The model included tumor surface area (TSA) to represent the number of tumor cells and microinvasion beyond the subserosa to represent bloodstream access for DNA fragments.

METHODS

- Blood samples were prospectively collected from participants with newly diagnosed untreated cancer and from participants without a diagnosis of cancer (Figure 1).
- Specimens in the first CCGA substudy were subjected to a prototype whole-genome bisulfite sequencing (WGBS) assay with a machine-learning classifier that predicted cancer or non-cancer. Samples in the second CCGA substudy were subjected to a TM assay with a newly trained machine-learning classifier that predicted cancer or non-cancer, with tissue of origin predicted for samples with cancer signal detected.1-3
- Tumor fraction was determined by comparing variants detected in a targeted sequencing assay of cfDNA from these participants to variants from a whole-genome sequencing (WGS) assay of matched, non-invasive formalin-fixed, paraffin-embedded tumor samples, while using a white blood cell WGS assay to control for germline variants.4
- To predict tumor fraction for non-metastatic colorectal adenocarcinomas, we constructed a simple physical model (TF = n * TSA) that assumes a linear relationship (slope, m) with tumor surface area (TSA) and interacts with microinvasion beyond or below the subserosa.
- The model was fit to 17 CRC participants (7 stage I, 5 stage II, and 5 stage III) from the first discovery substudy and was used to predict on 40 CRC participants (2 stage I, 11 stage II, and 13 stage III) from the training set from the second substudy with predictor variables available.

RESULTS

- The WGS assay classified CRC in 3 of 9 (33%) stage I cases, 11 of 18 (61%) stage II cases, and 17 of 30 (57%) stage III cases.
- The TM assay detected CRC in 6 of 13 (46%) stage I cases, 16 of 22 (73%) stage II cases, and 29 of 38 (76%) stage III cases.
- TF predicted CRC detection in the first CCGA substudy with an area under the curve of 0.948.
- For the WGBS assay, the classifier score for CRC cases increased approximately linearly with log(TF) before saturation for high tumor fractions (Figure 2).

Figure 1. Study Design and CONSORT Diagram

Figure 2. Detection of Cancer Increases With Increasing TF

The WGS assay assigns each classifier observation to one of three classes: cancer true (cancer TF >0.03), cancer false (cancer TF <0.03), or non-cancer (TF <0.03). Figure 3. TF Distribution by Clinical Stage (A) and Extent of Microinvasion (B)

- TF distributions overlapped for clinical stages I, II, and III (Figure 3A), while TF distributions were clearly separated by microscopic tumor extension (pathological T stage) alone (Figure 3B).
- Within a given extent of microinvasion (shallow or deep), TF can be modeled to increase linearly with TSA (Figures 4A and 5A).
- The model predicted m = 3.81×10−6/mm² for tumors that invaded beyond the subserosa (T stage IV; TF >0.03), which was 4.4 times higher than estimates for tumors below the subserosa (m = 0.866×10−6/mm²).

Figure 3. TF Distribution by Clinical Stage (A) and Extent of Microinvasion (B)

Figure 4. TF Model Including Both Tumor Size and Microinvasion in the First CCGA Substudy

- The model trained on the WGBS assay predicted CRC detection in the TM assay with an area under the curve of 0.844 (Figure 5).

Figure 5. Receiver Operating Characteristic Curve Based on Model TSA and Microinvasion in the Second CCGA Substudy

CONCLUSIONS

- Higher cfDNA TF was significantly associated with cancer detection.
- The strong relationship between cfDNA WGBS classifier scores and TF in the first substudy implies that TF might be impacted by training a classifier on cfDNA assay results.
- A model using TSA of tumor (or cancer) microinvasion (bloodstream access) predicted TF.
- Tumors not penetrating the subserosa had low cfDNA shedding into the bloodstream that likely limited detection.
- As a greater extent of tumor penetration may be associated with higher-risk diseases, these results suggest that the classifier may preferentially detect higher-risk cancers.
- Careful examination of pathologic features of CRC suggested that size (TSA for CRC) and access to circulation (microinvasion of CRC) of solid tumors may strongly affect the shedding of cfDNA into the bloodstream.

REFERENCES


Figure 4. TF Model Including Both Tumor Size and Microinvasion in the First CCGA Substudy

Figure 5. Receiver Operating Characteristic Curve Based on Model TSA and Microinvasion in the Second CCGA Substudy

TF model for the WGBS assay using only TSA and Microinvasion provides detection of CRC in the second substudy using the TM assay, AUC, and simultaneous determination of tissue of origin. cfDNA, circulating cell-free DNA; CRC, colorectal cancer; FPR, false positive rate; TF, tumor fraction; TPR, true positive rate for the prediction of cancer detection.