Early-stage cancer results in a multiplicative increase in cell-free DNA originating from healthy tissue

Konstantinos Mamis[1] and Ivana Bozic[1,2]

[1]*Department of Applied Mathematics, University of Washington, Seattle*
[2]*Herbold Computational Biology Program, Fred Hutchinson Cancer Center, Seattle*

Abstract

Sources of elevated cell-free DNA (cfDNA) in early-stage cancer patients are poorly understood. Leveraging a rich dataset of cfDNA in healthy individuals and cancer patients, we find that a large majority of cfDNA in cancer patients originates from non-tumor sources, and that the presence of early-stage cancer results in a multiplicative increase in the concentration of cfDNA originating from healthy tissue. The magnitude of the multiplicative increase is cancer-type specific and ranges from a ~1.3-fold increase in lung cancer, to a ~12-fold increase in patients with liver cancer. As cfDNA is cleared in the liver, the large increase in patients with liver cancer may imply that the systemic increase in cfDNA levels in the presence of cancer is due to slower clearance rate rather than higher cell turnover.

Main

Identifying biomarkers that can reliably detect cancer in early stages is one of the major challenges in oncology[1]. One promising biomarker is cell-free DNA (cfDNA), composed of the DNA fragments not bound to cells, that can be detected in the circulation. Increased levels of cfDNA are often reported in cancer patients[2-5], in particular in later stages[6]. However, the mechanisms by which the presence of cancer affects cfDNA are not fully understood[7–9].

Recent work demonstrates that the majority of cfDNA in cancer patients with highly elevated cfDNA originates from leukocytes, and that the presence of cancer likely results in a systemic increase in cfDNA concentration in these patients[10]. Whether a similar systemic increase due to cancer is present in all cancer patients is currently not known.

In this setting, several questions must be answered in order to better understand the mechanics of DNA circulating in plasma, which could lead to improving the accuracy of liquid biopsy screening tests. In what way, and how much does cancer presence affect the total cfDNA concentration? Is this effect cancer type-specific, and does it differ between early cancer stages? In the present work, we address these questions by employing a large dataset from Cohen *et al.*[11] for healthy individuals and patients with stage I-III ovarian, liver, stomach, pancreatic, esophageal, colorectal, lung, and breast cancer. This study[11] reports the cfDNA concentrations in plasma, as well as the concentration and mutant allele frequency (MAF) of DNA fragments in plasma harboring mutations in cancer driver genes (Methods); we refer to this portion of cfDNA as circulating mutant DNA (cmDNA).

We observe that the 631 healthy individuals from Cohen *et al.*[11] (Methods) can be divided into two groups: 572 (90.6%) that have cmDNA detected and have median cfDNA concentration 2.83ng/mL, and 59 (9.4%) that have no mutations detected in plasma and exhibit significantly (54 times) lower cfDNA concentrations (median 0.05ng/mL) (Fig. S1, Methods). Determining the reason why a portion of healthy population exhibits such low cfDNA concentrations requires further experimental investigation and is out of the scope of the present study. Since we aim to quantify and compare the effect of cancer on both cfDNA and cmDNA concentrations, we focus on the 90.6% of healthy individuals with cmDNA detected in plasma.

The mutant allele frequency for the most prevalent mutation detected in plasma of cancer patients of all types is low - median MAF is 0.06%, and 88.2% of mutations detected in

plasma have MAF below 1%. Combined with the fact that 82% of cancer patients' tumors harbor mutations (at high frequency) that can be detected by the liquid biopsy panel from Cohen *et al.*[11], this indicates that early-stage tumors in all cancer types analyzed do not contribute significantly to cfDNA concentration; most of cfDNA originates from non-tumor (healthy) tissue.

It has been reported that most of cfDNA circulating in plasma comes from hematopoiesis in healthy individuals[12–15], and some cancer patients[13]. Recent work by Mattox *et al.*[10] found that, in cases of healthy individuals and cancer patients with highly elevated cfDNA levels, the vast majority (76%) of cfDNA arises from leukocytes. They report that, in cancer patients with highly elevated cfDNA levels, the presence of cancer leads to a systemic increase in cfDNA concentration arising from non-tumor sources[10]. Inspired by this finding, we asked if a similar increase in cfDNA may be present in all cancer patients. We found that the presence of stage I-III lung, breast, colorectal, pancreatic, ovarian, esophageal, stomach and liver cancer leads to a multiplicative increase in cfDNA concentration compared to healthy individuals (Fig. 1). The multiplicative increase we uncovered is not simply an increase in cfDNA concentration - we find that the entire dataset of cfDNA concentrations is multiplied by a positive factor in patients with cancer compared to healthy individuals (Kolmogorov-Smirnov test at significance level $p = 0.05$, Methods). This effect can be seen as a horizontal "shift" or translation when the cumulative distribution function is plotted on a logarithmic x-scale (Fig. 1).

The multiplicative increase in cfDNA concentration is specific to each cancer type: 1.29 (value range not rejected by KS test at 5% significance level: 1.15-1.48) for lung cancer, 1.15 (0.9-1.43) for stage I breast cancer, 1.79 (1.45-1.99) for stage II/III breast cancer, 2.54 (2.22-2.63) for colorectal, 2.69 (2.16-3.93) for pancreatic, 4.3 (3.69-5.59) for ovarian, 5.24 (3.85-7.2) for esophageal, 5.66 (5.49-5.85) for stomach, and 12.63 (10.94-16.81) for liver cancer

(Fig. 1, Fig. S2, Table S1, Methods). We note that the multiplicative increase in cfDNA concentration in these cancer types originates from healthy tissue, as the large majority of cfDNA in stage I-III cancer patients is not shed from the tumor. There are no statistically significant differences in cfDNA concentration between stages I to III, for all cancer types except for breast cancer (Methods).

The uniform increase in cfDNA concentration observed in all evaluated cancer types could stem from two sources: either cfDNA shedding from non-tumor sources is systemically increased by a multiplicative factor in patients with cancer, or the presence of cancer results in a similar-fold lower clearance of cfDNA from circulation. There is evidence that an excess of dying cells can lead to slower cfDNA clearance, e.g. during vigorous exercise; the presence of cancer may also lead to an increase in dying cells, potentially slowing down the clearance of cfDNA. Cell-free DNA is typically cleared in spleen, kidney, and the liver[7]; it is of note that cfDNA concentration is highest in patients with liver cancer (~12 times increase in cfDNA concentration compared to healthy individuals), indicating the slower cfDNA clearance rate as a potential mechanism of how tumor presence results in the multiplicative increase of cfDNA levels.

We next hypothesized that, if cfDNA shed from non-tumor sources exhibits a multiplicative increase in the presence of cancer, then a similar increase should also be observed in circulating mutant DNA (cmDNA) that does not originate from tumor. In the dataset we use[11], both the most prevalent mutation in plasma and the mutations detected in tumor are reported for 645 of the 959 cancer patients (Methods). Comparing the cmDNA concentration in healthy individuals and in cancer patients in which cmDNA detected in plasma is not detected in tumor, we find uniform increases by a multiplicative factor for every cancer type: 1.92 (1.42-2.52) for lung, 2.42 (2.12-2.76) for breast, 2.36 (1.92-2.66) for colorectal, 4.01 (1.55-7.34) for pancreatic, 3.31 (1.76-5.38) for esophageal, 4.38 (2.74-5.77)

for stomach, and 12.37 (5.54-24.62) for liver cancer (Fig. 2, Fig. S3, Table S2, Methods). We observe that the ranges of multiplicative factor values for cfDNA and cmDNA from healthy tissue are overlapping for each of the cancer types (Fig. S4, Methods), except for breast cancer where ranges are comparable yet non-overlapping.

Mutant DNA in plasma may originate from healthy or tumor tissue; however, it may also be detected artificially due to sequencing errors. In a more recent work[16], Cohen *et al.* found that the artefactual mutations in plasma appeared with a MAF lower than 0.013%. Hence, we perform the cmDNA analysis again, disregarding the mutations with low MAFs (Table S3, Methods); the results remain similar to the results obtained for all mutations (Fig. S4, Methods).

The agreement between the multiplicative factors in total cfDNA concentration, and the concentration of cmDNA not originating from tumor, is a further validation that the presence of early-stage cancer results in a multiplicative shift towards higher concentrations of cfDNA (both wildtype and mutant) shed from healthy tissue; this increase is quantifiable and cancer type-specific. At this point, we cannot answer definitively whether this increase is the result of excessive cfDNA shedding or slower cfDNA clearance; it is also possible that these two causes are related: an initial, localized increase in cell turnover due to tumor presence could result in excess cfDNA that would slow down the cfDNA clearance rate from circulation.

Note that, the reported cfDNA concentrations may differ significantly, depending on the method used for the analysis of the samples. However, using additional data from Cohen *et al.*[17], we observe once again the effect of multiplicative factor between cfDNA concentrations from samples of healthy individuals that were analyzed using different methodologies (Fig. S5, Methods).

Last, the cancer-specific multiplicative factor may be affected by the effects of therapy or anesthesia. More specifically, we observe a reduced multiplicative factor for a cohort of

pancreatic cancer patients, many of whom had their blood sample taken after administration of anesthesia[17]. This was on par with the reported reduction in cfDNA concentration due to anesthesia[18] (Fig. S6, Methods).

Acknowledgements. The authors thank Bert Vogelstein for helpful comments on the manuscript. This work is supported by the National Science Foundation grant DMS-2045166. K.M. is also supported by the Pacific Institute for the Mathematical Sciences (PIMS).

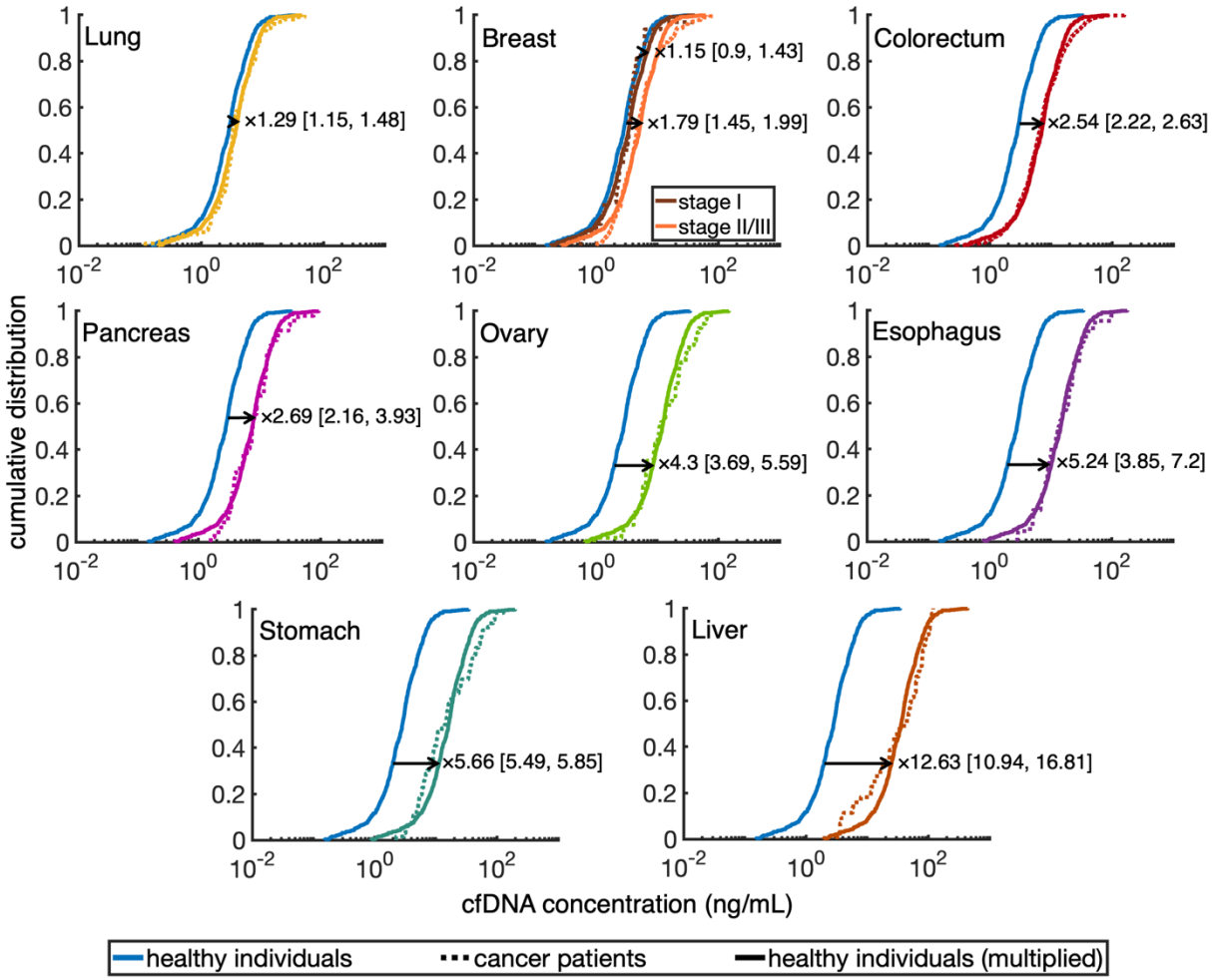

Figure 1: **Quantifying the multiplicative increase in cfDNA concentration in the presence of cancer.** Cumulative distribution functions for the cfDNA concentration in healthy individuals with mutations detected in plasma (N=572) and patients with lung (N=103), stage I breast (N=32), stage II/III breast (N=177), colorectal (N=385), pancreatic (N=47), ovarian (N=54), esophageal (N=44), stomach (N=66) and liver (N=44) cancer. Median empirical cfDNA concentrations (corresponding to 0.5 value of cumulative distribution) are 2.83 ng/mL for healthy individuals, and 3.31, 3.42, 4.60, 6.62, 7.76, 11.18, 13.97, 14.14, and 42.27 ng/mL for patients with lung, stage I breast, stage II/III breast, colorectal, pancreatic, ovarian, esophageal, stomach and liver cancer, respectively.

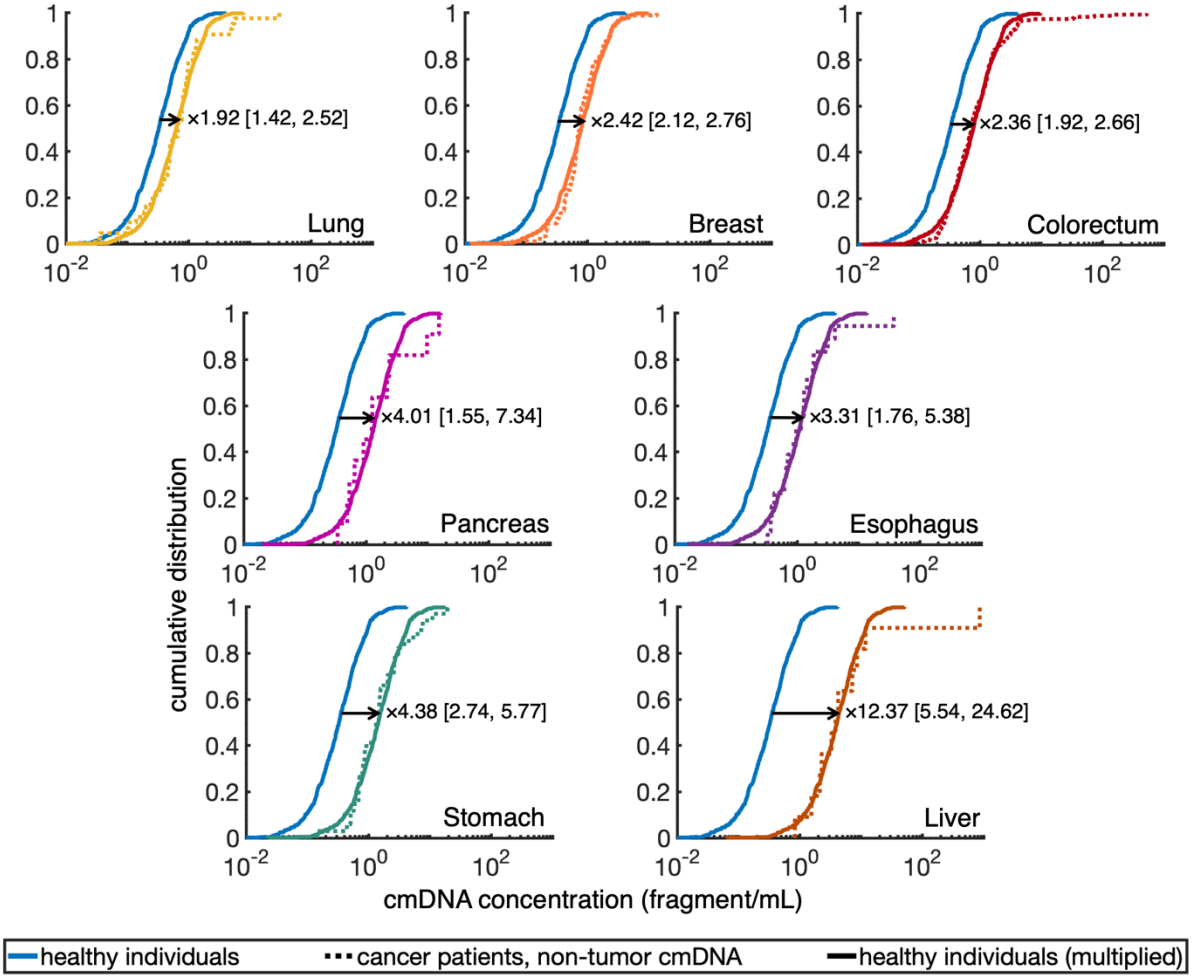

Figure 2: **Mutations detected in plasma of cancer patients but not detected in tumor exhibit a multiplicative increase in concentration compared to mutations detected in plasma of healthy individuals.** Cumulative distribution functions for the concentration of the most prevalent mutation found in plasma of healthy individuals (N=572) and patients with lung (N=43), breast (N=90), colorectal (N=205), pancreatic (N=11), esophageal (N=18), stomach (N=34) and liver (N=11) cancer in which the most prevalent mutation in plasma was not detected in tumor. Median empirical cmDNA concentrations (corresponding to 0.5 value of cumulative distribution) are 0.33 fragment/mL for healthy individuals, and 0.67, 0.73, 0.72, 1.23, 1.05, 1.39, and 3.85 fragment/mL for patients with cancer of the lung, breast, colorectum, pancreas, esophagus, stomach and liver, respectively.

References

1. Crosby, D. *et al.* Early detection of cancer. *Science (1979)* **375**, (2022).

2. Alese, O. B. *et al.* Circulating Tumor DNA: An Emerging Tool in Gastrointestinal Cancers. *American Society of Clinical Oncology Educational Book* **42**, 279–298 (2022).

3. Bronkhorst, A. J., Ungerer, V. & Holdenrieder, S. The emerging role of cell-free DNA as a molecular marker for cancer management. *Biomol Detect Quantif* **17**, (2019).

4. Diehl, F. *et al.* Circulating mutant DNA to assess tumor dynamics. *Nat Med* **14**, 985–990 (2008).

5. Roncucci, L. *et al.* The influence of age on colonic epithelial cell proliferation. *Cancer* **62**, 2373–2377 (1988).

6. Bettegowda, C. *et al.* Detection of Circulating Tumor DNA in Early- and Late-Stage Human Malignancies. *Sci Transl Med* **6**, (2014).

7. Kustanovich, A., Schwartz, R., Peretz, T. & Grinshpun, A. Life and death of circulating cell-free DNA. *Cancer Biol Ther* **20**, 1057–1067 (2019).

8. Siravegna, G., Marsoni, S., Siena, S. & Bardelli, A. Integrating liquid biopsies into the management of cancer. *Nat Rev Clin Oncol* **14**, 531–548 (2017).

9. Diehl, F. *et al.* Detection and quantification of mutations in the plasma of patients with colorectal tumors. *Proceedings of the National Academy of Sciences* **102**, 16368–16373 (2005).

10. Mattox, A. K. *et al.* The Origin of Highly Elevated Cell-Free DNA in Healthy Individuals and Patients with Pancreatic, Colorectal, Lung, or Ovarian Cancer. *Cancer Discov* **13**, 2166–2179 (2023).

11. Cohen, J. D. *et al.* Detection and localization of surgically resectable cancers with a multi-analyte blood test. *Science (1979)* **359**, 926–930 (2018).

12. Snyder, M. W., Kircher, M., Hill, A. J., Daza, R. M. & Shendure, J. Cell-free DNA Comprises an in Vivo Nucleosome Footprint that Informs Its Tissues-Of-Origin. *Cell* **164**, 57–68 (2016).

13. Sun, K. *et al.* Plasma DNA tissue mapping by genome-wide methylation sequencing for noninvasive prenatal, cancer, and transplantation assessments. *Proc Natl Acad Sci U S A* **112**, E5503–E5512 (2015).

14. Zheng, Y. W. L. *et al.* Nonhematopoietically derived DNA is shorter than hematopoietically derived DNA in plasma: A transplantation model. *Clin Chem* **58**, 549–558 (2012).

15. Heitzer, E., Auinger, L. & Speicher, M. R. Cell-Free DNA and Apoptosis: How Dead Cells Inform About the Living. *Trends Mol Med* **26**, 519–528 (2020).

16. Cohen, J. D. *et al.* Detection of low-frequency DNA variants by targeted sequencing of the Watson and Crick strands. *Nat Biotechnol* **39**, 1220–1227 (2021).

17. Cohen, J. D. *et al.* Combined circulating tumor DNA and protein biomarker-based liquid biopsy for the earlier detection of pancreatic cancers. *Proc Natl Acad Sci U S A* **114**, 10202–10207 (2017).

18. Prakash, K., Aggarwal, S., Bhardwaj, S., Ramakrishna, G. & Pandey, C. K. Serial perioperative cell-free DNA levels in donors and recipients undergoing living donor liver transplantation. *Acta Anaesthesiol Scand* **61**, 1084–1094 (2017).

Methods

Patient data

Cohen *et al.*[11] report the concentration (ng/mL) of total cell-free DNA (cfDNA) detected in plasma for 812 healthy individuals and for: 104 patients with lung cancer, 209 with breast cancer, 388 with colorectal cancer (CRC), 93 with pancreatic cancer, 54 with ovarian cancer, 45 with esophageal cancer, 68 with stomach cancer, and 44 with liver cancer. All cancer cases reported are in stages I to III. The dataset includes samples from 181 healthy individuals and 46 pancreatic cancer evaluated in a previous study[17]. In order to ensure that all samples were analyzed following the same process, we exclude these samples from the present work.

Cohen *et al.*[11] also report the concentration (fragment/mL) of the circulating mutant DNA (cmDNA) of the most prevalent DNA mutation detected in plasma for the same population of healthy individuals and cancer patients. The panel for mutation detection from plasma includes 61 amplicons, each querying an average of 33 base pairs within one of the 16 cancer driver genes. Genes included in the panel and the number of amplicons for each gene are: *NRAS* (2), *CTNNB1* (2), *PIK3CA* (4), *FBXW7* (4), *APC* (2), *EGFR* (1), *BRAF* (1), *CDKN2* (2), *PTEN* (4), *FGFR2* (1), *HRAS* (1), *KRAS* (3), *AKT1* (1), *TP53* (31), *PPP2R1A* (1), *GNAS* (1).

Apart from the most prevalent mutation in plasma, mutations found in tumor are also reported for 336 of the patients with colorectal cancer; the respective number of patients is 70 for lung cancer, 116 for breast cancer, 15 for pancreatic cancer, 7 for ovarian cancer, 36 for esophageal cancer, 47 for stomach cancer, and 24 for liver cancer. Since it is important to compare mutations detected in the tumor and plasma of the same patient, we focus our analysis regarding cmDNA at the aforementioned subset of patients with known mutations in both plasma and tumor; we also exclude the 1 lung cancer patient, 2 CRC patients, 1

esophageal cancer patient, and 2 stomach cancer patients, for which no cmDNA was detected in plasma. Due to the small dataset (7 patients), we exclude ovarian cancer from the cmDNA analysis.

Kolmogorov-Smirnov test

Kolmogorov-Smirnov (KS) test is a nonparametric test for continuous one-dimensional probability distributions that we employ to check whether two datasets may come from the same distribution, or whether a dataset is drawn from a particular distribution[19]. More specifically, in the case of two datasets, we use the Kolmogorov-Smirnov test to check whether the multiplication of all elements of one dataset by the same deterministic factor may result in the two datasets originating from the same distribution.

Difference in cfDNA/cmDNA concentrations between different cancer stages

Cohen *et al.*[11] report the cfDNA concentrations for patients in stages I/II/III of various cancer types. The number of patients in each stage is 46/27/31 for lung cancer, 32/114/63 for breast cancer, 77/191/120 for colorectal cancer, 2/39/6 for pancreatic cancer, 9/4/41 for ovarian cancer, 5/29/11 for esophageal cancer, 21/30/17 for stomach cancer and 5/19/20 for liver cancer. By performing KS tests between the samples of cfDNA concentrations in different stages of the same cancer type, the hypothesis that the samples originate from the same distribution is not rejected, at a 5% significance level, for all cases except for breast and liver cancer: KS test p-value between samples for stages II and III breast cancer is 0.59, p-values between stage I and II, and stage I and III breast cancer are 0.005 and 0.001 respectively. Thus, cfDNA concentrations in stage I breast cancer are significantly different than the concentrations in stages II/III. The same is observed for liver cancer stages, with p-values

0.01, <0.001, and 0.18 for the KS tests between stages I and II, I and III, and II and III respectively; however, for liver cancer, the sample for stage I patients is small (5 patients). For cmDNA samples from cancer patients and mutation not detected in tumor, we observe no significant differences in concentrations between sexes or stages.

Multiplicative factors for increased cfDNA/cmDNA in cancer patients

We determine the value of multiplicative factor a for which we have the minimum Kolmogorov distance between the empirical CDF for a times the sample of cfDNA concentration in healthy patients with mutations detected in plasma, and the empirical CDF for the sample of cfDNA concentration in patients of each cancer type. We report the value of the multiplicative factor a, and the p-value of KS test between the sample of cfDNA concentration in healthy patients with mutations detected in plasma multiplied by a, and the cfDNA concentration in patients of each cancer type (Table S1).

We also report the value range of the multiplicative factor a, for which the KS test does not reject, at 5% significance level, the hypothesis that cfDNA data for cancer patients and the cfDNA data for healthy patients multiplied by the respective factor come from the same distribution. We also observe the agreement between the datasets in probability-probability (P-P) plots (Fig. S2), which stay close to the diagonal for all types of cancer.

We also calculate the multiplicative factors a between the cmDNA concentrations reported in healthy individuals, and the concentrations of cmDNA not shed from tumor in patients of the various cancer types. The value of a resulting in the least Kolmogorov distance between the empirical CDFs, along with the 5% significance level ranges, are shown in Table S2; the respective P-P plots, which also stay close to the diagonal, are shown in Fig. S3. We exclude ovarian cancer from cmDNA concentration analysis, due to insufficient data: there are only 7 ovarian cancer patients with cmDNA reported. We also calculate the multiplicative factors

for cmDNA concentrations, excluding the samples with MAF less than i) 0.01%, ii) 0.015%, and iii) 0.02%, since mutations detected in such low MAFs may be artefactual[16]. We report the numbers healthy individuals and cancer patients with cmDNA not from tumor for different cutoff values for MAF in Table S3. The values of multiplicative factors are similar for all cases considered (Table S2).

Dependence of cfDNA concentrations on the sample analysis methodology

The reported cfDNA concentrations depend on the method used for the analysis of the samples. In a 2017 study[17], Cohen *et al.* analyzed samples from 182 healthy individuals using different methodology, and reported significantly higher cfDNA concentrations (median 4.25ng/mL, Kolmogorov-Smirnov test p-value < 0.001) than in the dataset we use (Cohen et al. 2018)[11]. The cfDNA concentrations in healthy individuals in the two studies differ by the multiplicative factor 1.57 (1.34-1.79); meaning the distributions for cfDNA concentrations are similar in shape, although shifted. (Fig. S5).

Effect of anesthesia on the multiplicative factor

The cancer-specific multiplicative factor may be affected by the effects of therapy or anesthesia. Anesthesia has been reported to reduce the concentration of cfDNA ~2.5-fold[18]. Cohen *et al.* (2017)[17] report cfDNA concentrations from 182 healthy individuals and 221 patients with stage I-III pancreatic cancer. As many pancreatic cancer patients in this study had their blood sample taken after administration of anesthesia, we expect that the multiplicative factor for pancreatic cancer is up to 2.5-fold lower compared to Cohen *et al.* (2018)[11], which excluded patients where anesthesia was known to be administered prior to sample collection. Indeed, we find that the multiplicative factor for pancreatic cancer from Cohen *et al.* (2017) [17] is 1.45 (1.30-1.59) (Fig. S6). This factor range is ~2 to ~2.5 times lower that the multiplicative factor for pancreatic cancer in Cohen *et al.* (2018) [11] data, as expected.

Additional References

19. NIST/SEMATECH. e-Handbook of Statistical Methods, Section 1.3.5.16 http://www.itl.nist.gov/div898/handbook/. (2023).

Cancer	multiplicative factor	KS *p*-value
Lung	1.29 [1.15, 1.48]	0.62
Stage I Breast	1.15 [0.90, 1.43]	0.78
Stage II/III Breast	1.79 [1.45, 1.99]	0.48
Colorectum	2.54 [2.22, 2.63]	0.30
Pancreas	2.69 [2.16, 3.93]	0.82
Ovary	4.30 [3.69, 5.59]	0.44
Esophagus	5.24 [3.85, 7.20]	0.94
Stomach	5.66 [5.49, 5.85]	0.10
Liver	12.63 [10.94, 16.81]	0.28

Table S1: Multiplicative factors for cfDNA concentration, for each cancer type. We report the *p*-value of the KS test between cfDNA samples from patients of each cancer and the cfDNA samples from healthy individuals with cmDNA detected multiplied by the corresponding multiplicative factor. We also report the range of values for the multiplicative factor with KS test *p*-values >0.05.

Cancer	multiplicative factor	KS *p*-value	multiplicative factor MAF >0.01%	KS *p*-value	multiplicative factor MAF >0.015%	KS *p*-value	multiplicative factor MAF >0.02%	KS *p*-value
Lung	1.92 [1.42, 2.52]	0.89	1.93 [1.45, 2.52]	0.88	1.89 [1.41, 2.45]	0.88	1.75 [1.34, 2.34]	0.89
Breast	2.42 [2.12, 2.76]	0.41	2.50 [2.22, 2.79]	0.27	2.44 [2.16, 2.66]	0.31	2.35 [2.06, 2.54]	0.31
Colorectum	2.36 [1.92, 2.66]	0.51	2.49 [1.98, 2.70]	0.42	2.51 [2.00, 2.74]	0.45	2.59 [2.02, 2.89]	0.44
Pancreas	4.01 [1.55, 7.34]	0.86	4.01 [1.43, 8.53]	0.82	4.01 [1.40, 8.47]	0.82	4.72 [1.39, 9.35]	0.78
Esophagus	3.31 [1.76, 5.38]	0.99	3.36 [1.77, 6.07]	0.99	3.64 [1.74, 6.57]	0.99	3.47 [1.52, 6.78]	0.99
Stomach	4.38 [2.74, 5.77]	0.86	4.56 [2.99, 6.25]	0.87	5.19 [2.98, 7.70]	0.83	5.01 [2.70, 6.71]	0.62
Liver	12.37 [5.54, 24.62]	0.99	14.56 [6.88, 37.14]	0.99	17.78 [7.19, 46.06]	0.99	16.96 [7.03, 43.43]	0.99

Table S2: Multiplicative factors for cmDNA concentration not shed from tumor, for each cancer type. We consider all samples, and samples with MAF >0.01%, 0.015%, 0.02%, and report the *p*-value of the KS test between cmDNA samples from patients of each cancer and the cmDNA samples from healthy individuals multiplied by the corresponding multiplicative factor. We also report the range of values for the multiplicative factor with KS test *p*-values >0.05.

	Samples no. all MAF	Samples no. MAF >0.01%	Samples no. MAF >0.015%	Samples no. MAF >0.02%
Healthy	572	561	539	496
Lung	43	42	42	42
Breast	90	87	85	82
Colorectum	205	186	174	156
Pancreas	11	10	10	9
Esophagus	18	15	13	12
Stomach	34	31	25	22
Liver	11	9	7	7

Table S3: number of samples of healthy individuals with cmDNA detected, and cancer patients with cmDNA not shed from tumor, for various cutoff values for MAF.

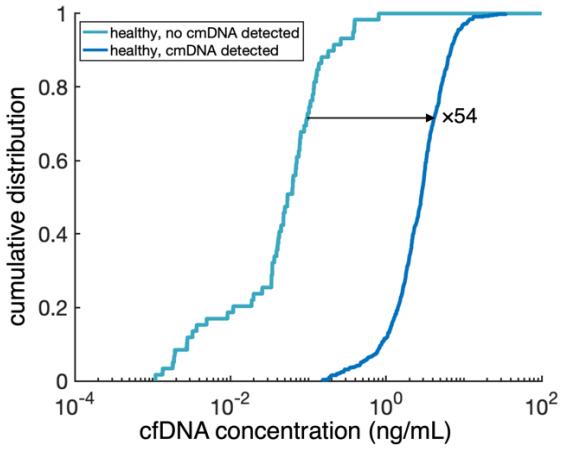

Figure S1: **Multiplicative factor effect in cfDNA concentration between healthy with or without cmDNA detected.** Empirical cumulative distribution functions for the cfDNA concentration in healthy individuals with no mutations detected in liquid biopsies (9.4% of healthy individuals) and healthy individuals with mutations detected in plasma (90.6% of healthy individuals). By performing the KS test, we observe that cfDNA concentrations in healthy individuals with no mutations detected are consistent with the cfDNA concentrations in healthy individuals with mutations detected in plasma, multiplied by 1/54 (*p*-value: 0.19).

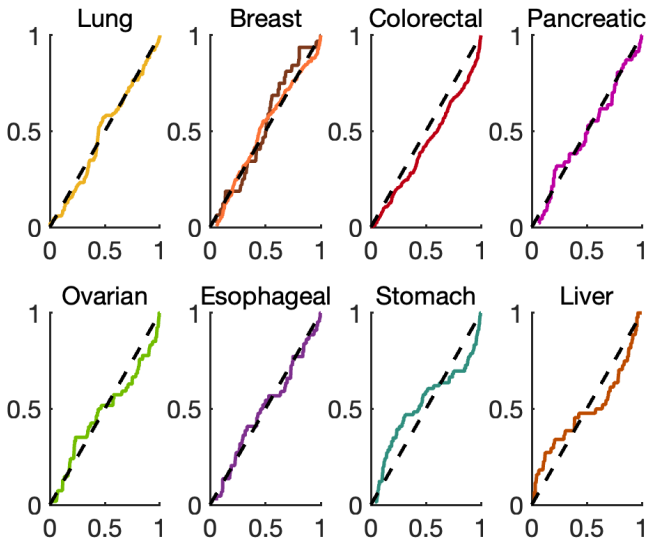

Figure S2: **P-P plots for the multiplicative factor effect in cfDNA concentration between healthy and cancer patients.** In each panel, we depict the probability pairs calculated from the empirical CDF for cfDNA concentration of cancer patients of a particular type, and the empirical CDF of cfDNA concentration in healthy patients multiplied by the respective factor from Table S1.

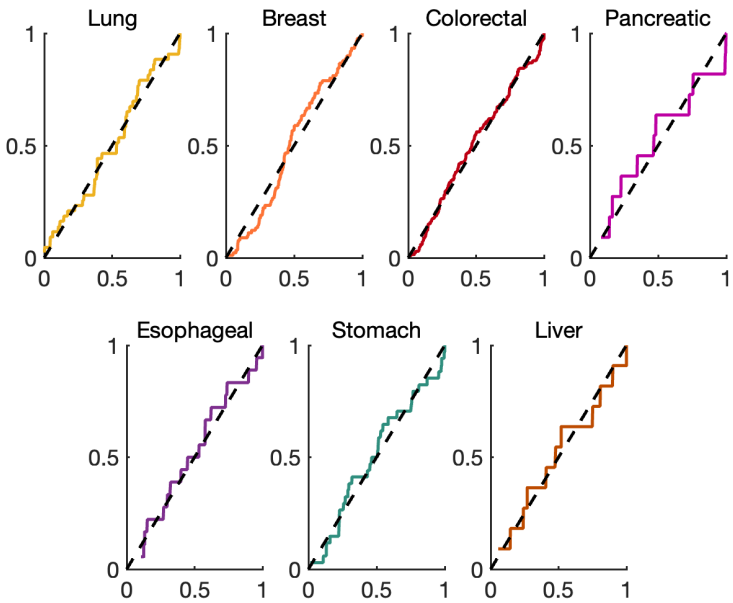

Figure S3: **P-P plots for the multiplicative factor effect in cmDNA concentration between healthy and cancer patients.** In each panel, we depict the probability pairs calculated from the empirical CDF for cmDNA concentration shed from non-tumor tissue of cancer patients of a particular type, and the empirical CDF of cmDNA concentration in healthy patients multiplied by the respective factor from Table S2 (values for all MAF). Ovarian cancer is not depicted due to insufficient data.

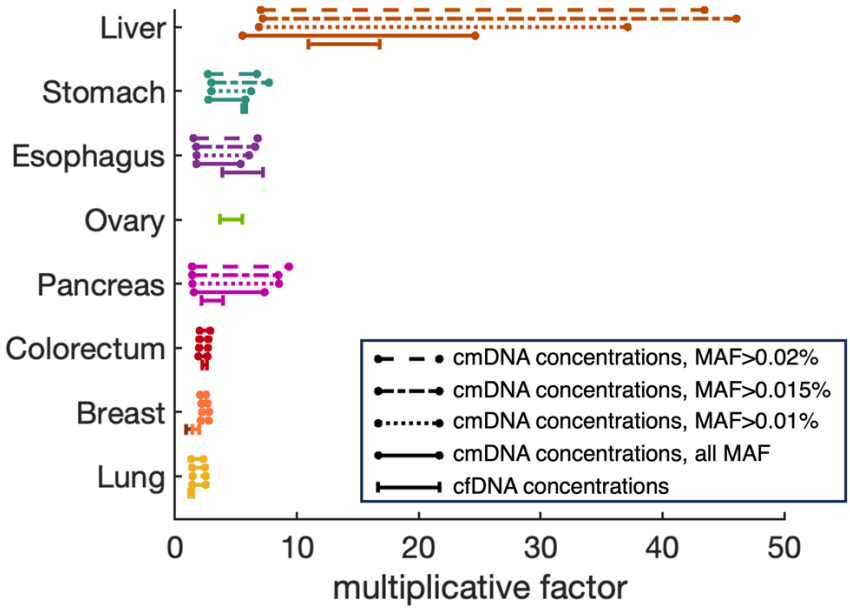

Figure S4: **Multiplicative factor ranges for each cancer type, quantifying the increase in concentrations of total cfDNA in plasma, and of cmDNA not shed from cancer tissue.** For each cancer type, we depict the ranges of multiplicative factor for cfDNA and cmDNA concentration from Tables S1, S2.

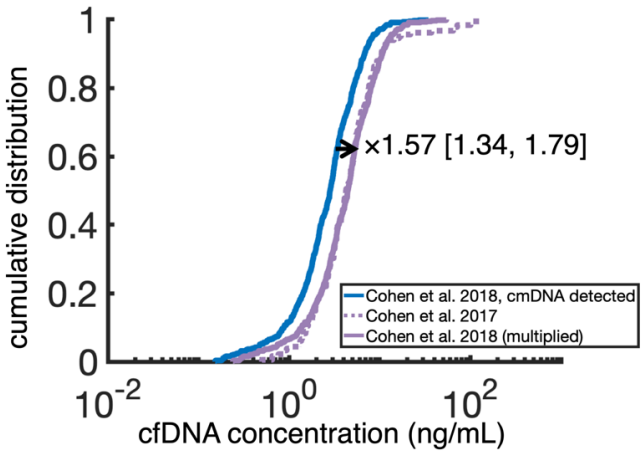

Figure S5: **Multiplicative factor effect in cfDNA concentration between healthy individuals reported between two studies[11, 17] by Cohen *et al.*** Empirical cumulative distribution functions for the cfDNA concentration in healthy individuals with no mutations detected from the 2018 study by Cohen *et al.*[11], and healthy individuals from the 2017 study by Cohen *et al.*[17]. By performing the KS test, we observe that cfDNA concentrations in healthy individuals from the 2017 study are consistent with the cfDNA concentrations in healthy individuals from the 2018 study, multiplied by 1.57 (p-value: 0.95).

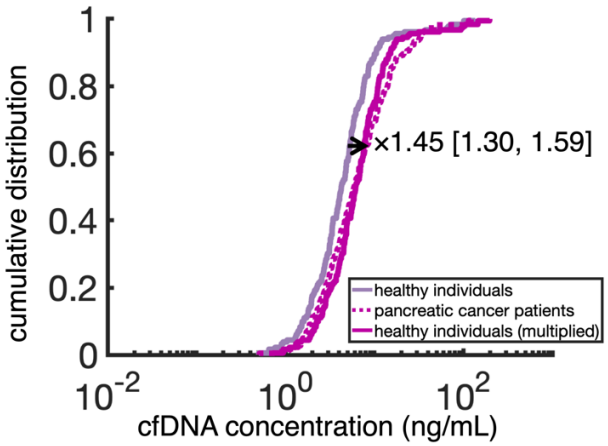

Figure S6: **Multiplicative factor effect in cfDNA concentration between healthy individuals and pancreatic cancer patients reported in Cohen *et al.*[17]** Empirical cumulative distribution functions for the cfDNA concentration in 182 healthy individuals, and 221 pancreatic cancer patients from the 2017 study by Cohen *et al.*[17]. By performing the KS test, we observe that cfDNA concentrations in healthy individuals are consistent with the cfDNA concentrations in pancreatic cancer patients, multiplied by 1.45 (*p*-value: 0.39).