Pitfalls in Cancer Biomarker Discovery and Validation with Emphasis on Circulating Tumor DNA Ⓐ Ⓒ

Annie H. Ren[1,2], Clare A. Fiala[1], Eleftherios P. Diamandis[1,2,3], and Vathany Kulasingam[2,3]

ABSTRACT

Despite significant investment of funds and resources, few new cancer biomarkers have been introduced to the clinic in the last few decades. Although many candidates produce promising results in the laboratory, deficiencies in sensitivity, specificity, and predictive value make them less than desirable in a patient setting. This review will analyze these challenges in detail as well as discuss false discovery, problems with reproducibility, and tumor heterogeneity. Circulating tumor DNA (ctDNA), an emerging cancer biomarker, is also analyzed, particularly in the contexts of assay specificity, sensitivity, fragmentation, lead time, mutant allele fraction, and clinical relevance. Emerging artificial intelligence technologies will likely be valuable tools in maximizing the clinical utility of ctDNA which is often found in very small quantities in patients with early-stage tumors. Finally, the implications of challenging false discoveries are examined and some insights about improving cancer biomarker discovery are provided.

See all articles in this *CEBP Focus* section, "NCI Early Detection Research Network: Making Cancer Detection Possible."

Introduction

A longstanding dream of clinical biochemists and medical oncologists is to use simple blood-based tests to diagnose cancer early. Such a development would allow for more effective, targeted, and less toxic treatments and would constitute a major win in the war against cancer (1). On the basis of this premise, the U.S. NCI (Bethesda, MD) formed the Early Detection Research Network (EDRN), an organization with a mandate to discover and validate cancer biomarkers (2). EDRN's long-term goal is to bring biomarkers to the clinic, so that they can contribute to better patient care through disease prevention and personalized management.

In recent years, technological advances such as genomic, proteomic, and other omic profiling of tumor tissues and fluids have helped to elucidate multifaceted genetic drivers in cancer, revealed cancer heterogeneity, and provided clues for cancer evolution. These advances have highlighted the need for a new generation of biomarkers that capture the complicated and ever evolving proteogenomic landscape of tumors. Although many new technologies have been employed in this effort (3, 4), this research area has been met with limited success. We have not yet been able to bridge the age-old gap between initial biomarker discovery and subsequent clinical translation (5–8). The reality is that the current handful of clinically useful circulating cancer biomarkers were discovered 40 to 60 years ago (9).

The task of discovering new and improved noninvasive biomarkers for cancer, under the general umbrella of "liquid biopsy" (i.e., analysis performed in a tube of blood) has likely been underestimated. In this review, we will attempt to explain some reasons for these apparent failures, despite an abundance of funding, the use of revolutionary multi-omics technologies, the availability of biobanks, and the utilization of innovative bioinformatics solutions. We will focus on circulating cancer biomarkers. The field of molecular biomarkers, which are usually analyzed in tumor tissue for prognosis, selection of therapy, and predicting therapeutic response, is beyond the scope of this review. Because many difficulties in biomarker discovery and validation have already been discussed by our group and others (10, 11), this review will place special emphasis on circulating tumor DNA (ctDNA), the most promising new cancer biomarker (12–14). To drive positive changes in cancer biomarker research, we will also delve into the efforts of our group and others to promote scientific debates, with the idea that some current obstacles could be overcome with lessons learned from the past.

Challenges with Biomarker Discovery

A literature search on cancer biomarkers will return thousands of hits. Why do many biomarkers that display promising *in vitro* results not make it to the clinic? There are two main reasons (15). One is that the performance of the majority of newly discovered biomarkers, as exemplified by their sensitivity and specificity, is often statistically significant ($P < 0.05$) when comparing study groups with benign lesions of similar tissue type versus malignant, cancerous tumors (15). Nevertheless, the new biomarkers are often clinically unhelpful. In other words, if the information provided by the biomarkers is either weak or not essential for clinical decision-making (nonactionable), clinicians will not use them. Second, even biomarkers that attempt to address a clear clinical unmet need, and show reasonable sensitivity and specificity, may not be adopted clinically if their positive predictive value (PPV) and negative predictive value (NPV) are below a clinically defined threshold. It is not immediately apparent to the nonspecialist, but the PPV of a test (the chances for having a disease if the test is positive) is dependent not only on the sensitivity and specificity of the test but also on the disease prevalence among the tested population. For relatively rare cancers (such as pancreatic and ovarian cancer, for which their prevalence in the tested population is <0.05%), a good screening test must be characterized by very high sensitivity (e.g., >90%) and even higher specificity (>99%) to yield a PPV of >10%. Furthermore, proving the clinical value of a new test requires detailed validation with large studies, which are expensive and time consuming (9). Consequently, diagnostic companies will be reluctant to

[1]Department of Pathology and Laboratory Medicine, Mount Sinai Hospital, Toronto, Ontario, Canada. [2]Department of Laboratory Medicine and Pathobiology, University of Toronto, Toronto, Ontario, Canada. [3]Department of Clinical Biochemistry, University Health Network, Toronto, Ontario, Canada.

Corresponding Author: Vathany Kulasingam, University of Toronto, 200 Elizabeth Street, Room 3EB 362A, Toronto M5G 2C4, Canada. Phone: 416-340-4800, ext. 8589; Fax: 416-340-4215; E-mail: vathany.kulasingam@uhn.ca

license, validate, manufacture, and sell a test that has questionable clinical indications. In short, the qualification bar for tests aiming to migrate from discovery to clinical use is very high. This is especially true for tests designed for cancer screening programs. For more detailed descriptions on the characteristics of diagnostic tests and the requirements for their clinical use, please see ref. 15.

False discovery and scientific reproducibility

Beyond the complex logistics of introducing a new biomarker into clinical practice, there are often pitfalls in the initial discovery phase that threaten the reliability and utility of what could first appear to be an auspicious new marker. Upon closer examination, most candidate biomarkers identified from preliminary studies fail validation in subsequent, more rigorous studies. False discovery (defined as a discovery that was made by an error) occurs when an initial finding related to a new biomarker cannot be reproduced in independent sets of samples or by other laboratories. The main reasons for false discovery are study design and execution deficiencies/errors, including preanalytical, analytical, postanalytical, and bioinformatic shortcomings, as we discussed previously (5, 15). For example, poor-quality clinical samples can contribute to false discovery, even when using a meticulous study design (16). Bias can often play a significant role in false biomarker discovery (17). Bias can be defined as a systematic error due to an unfair sampling (i.e., site to site variations and confounding variables such as age that were not taken into account in study design and data analysis), or poor execution of an experiment (17).

False discovery is closely related to the contemporary and highly discussed problem of scientific irreproducibility (18–21). Some limitations that lead to false discovery and irreproducibility are very subtle and can only be identified years after publication of seemingly revolutionary data. Examples of this sort have been described by us elsewhere (5, 15). Raising awareness on false discovery and scientific irreproducibility are critical first steps in reducing some dangers and enhancing the quality in biomarker research. We recently proposed some ideas to reduce scientific irreproducibility and false discovery (18, 20, 21). In **Table 1**, we summarize some general suggestions which could minimize false discovery in biomarker studies.

Tumor heterogeneity

The intricate nature of tumor biology is also to blame for the frequent failures in cancer biomarker research. In the past, cancer

Table 1. The do's and don'ts of biomarker discovery efforts.

1. Never use low-quality samples in the discovery phase.
2. Avoid bias between patients (cases) and nonpatients (control) groups.
3. Use reliable and trusted analytical techniques; if possible, verify results with an orthogonal (independent) technique.
4. Use appropriate statistical analyses and transparent bioinformatic tools (not black boxes).
5. Use enough samples to achieve statistical power.
6. Always validate discovery with one or more independent sets of samples, preferably from different institutions (ensure data generalizability).
7. Have a clear path to clinical decision-making after the test is performed in patients (actionability).
8. Check whether data make biological sense or are supported by previous literature.
9. Share data with one or more knowledgeable colleagues and be open to criticism; this could save you from future embarrassments.
10. Before publishing, file for possible intellectual property and present the data at a specialty scientific meeting, seeking input from participants.

was considered to be a single disease. With the surge in various genomic methods such as whole-exome and whole-genome sequencing, and even more recently single-cell sequencing, we have learned tremendously about the vast landscape of tumor heterogeneity in recent years. The diversity of genetic mutations is markedly palpable between tumors of the same organ, within the primary tumor, and between the primary and metastatic tumors (22). Innovations in imaging and computer-vision technologies have allowed for the spatial mapping of cancer cells in histologic samples to visualize heterogeneity between and within patients (23). Aside from intertumor and intratumor variability, tumors evolve over time, resulting in temporal heterogeneity that often accounts for chemoresistance, metastasis, and recurrence (24, 25). More recently, numerous studies have demonstrated the presence of cancer stem cells that drive mutations and create vastly different genetic intertumor and intratumor mutational landscapes over time (26). Mounting experimental data has espoused the coexistence of genetically distinct cell clones in tumors, which drive clonal evolution giving rise to the heterogeneity (27). Revelations in tumor heterogeneity can be harnessed to our advantage for translational utility. Studies have begun to focus on combining molecular subtyping, genomic profiling, and clinical data with traditional biomarkers to predict clinical outcomes in different cancers (28–30). New studies have also examined the association of the extent of heterogeneity and molecular subtypes in tumors with drug resistance, recurrence, and metastasis (30). Along these lines, molecular characterization of tumors has also been exploited for customizing therapy selection and patient management (31–34). However, this new knowledge also suggests that it is highly unlikely we will find circulating cancer biomarkers that can be equally informative in all, or most, patients. More likely, tumor heterogeneity gives rise to biomarkers that are related to the mutational landscapes of each patient (see also below).

Personalized biomarkers

With increasing evidence on the high variability in tumor composition among patients with the same cancer type, there is a new sense of urgency to challenge the status quo of biomarker research. We propose to slowly shift from the orthodox discovery of biomarkers that are common for a cancer type to screening of serum samples of each patient, identifying the most informative set of markers for their unique tumor. We recently suggested this concept and coined the putative markers as "personalized" or "rare" tumor markers (35). We hypothesized, based on the status of the biomarker field over the past decades, that it is unlikely to unveil new biomarkers that perform with high specificity and sensitivity for all patients, and surpass the performance of already established and widely used biomarkers (such as PSA for prostate cancer or CA125 for ovarian cancer). There are hundreds to thousands of biomarkers in the literature that show high specificity but low sensitivity and these have not been further studied as a result (35). Nevertheless, compiling a library of such "rare" tumor markers, which are informative in only 5%–15% of patients, may offer a robust tool to screen patients to identify a custom set of markers that are most informative for their unique tumor (35). This approach could capture the heterogeneity of each tumor. In the future, personalized biomarkers, tailored to each patient, could find applicability in selecting targeted therapies, determining the extent of therapeutic response and monitoring for signs of recurrence.

ctDNA

The idea of finding noninvasive, circulating biomarkers that provide comprehensive information on the genetic makeup of each tumor is

Table 2. Some current methods to diagnose cancer based on cell-free DNA analysis.

Molecular analysis of ctDNA/cfDNA	Genetic alterations	Sequencing depth	Control	Reference
Targeted sequencing of 507 genes previously associated with cancer	Single-nucleotide variants, insertions, and deletions	60,000	WBC DNA analysis	(38)
Whole-genome sequencing	Copy-number variations	30	WBC DNA analysis	(39, 40)
Whole-genome bisulfite sequencing	Methylation differences	30	WBC DNA analysis	(38, 42)
Genome-wide fragmentation patterns of cfDNA	Altered fragmentation profiles in a position-specific manner	0.5–2	cfDNA from healthy individuals	(14, 41)

Abbreviation: WBC, white blood cell.

gaining traction. In the last few years, the emergence of genome sequencing, bolstered by the analysis of cell-free or circulating-free DNA (cfDNA) in the circulation, created considerable excitement. The potential is considerable for noninvasive cfDNA testing as one of the most promising new cancer biomarkers, particularly to pave the way toward personalized biomarker tests (36–42). cfDNA in the circulation can be derived from both normal and tumor cells, with the latter fraction known as ctDNA. By isolating and sequencing cfDNA using various genomic techniques, the small fraction of ctDNA can be detected and distinguished from cfDNA originating from normal cells based on genetic alterations, epigenetic modifications, or fragment length (**Table 2**). This distinction forms the basis for noninvasive pan-cancer detection tests, where blood-based samples can be screened for ctDNA in individual patients to assess tumor burden, as a form of liquid biopsy.

Numerous academic and commercial laboratories have focused on ctDNA testing and have developed sophisticated technologies for detecting the common genetic abnormalities associated with ctDNA, with impressive preliminary clinical results. The sensitivity of cancer detection, depending on cancer type, ranges from 50% to 70% at 90%–98% specificity. Recent revelations from CancerSEEK reported by Cohen and colleagues showed that combining information of genetic changes found in ctDNA with information from traditional circulating cancer biomarkers increases specificity to >99% and sensitivity up to 98% for some cancers (40). The advent of ctDNA holds potential for various important applications, including early diagnosis, prognosis, monitoring therapy success, and estimating tumor volume (43). Although these applications are now being slowly implemented in patient care, early detection of cancer using ctDNA markers has not yet reached the clinic. This underwhelming clinical application of ctDNA testing can be attributed to the general challenges mentioned above for circulating protein biomarkers. However, in this study, we wish to draw attention to some key issues that are particular to ctDNA as a cancer biomarker, with emphasis on early detection.

Assay sensitivity

So far, all published investigations regarding cancer detection using ctDNA have reported sensitivity that was calculated from retrospectively collected samples with a mixture of patients with early- and late-stage cancers (40–42). These cancers were invariably diagnosed by clinical symptoms and imaging/biopsy. Thus, to better understand the ability of ctDNA tests to detect cancer preclinically, at an early stage, we previously used published empirical data to correlate tumor size with the expected amount of ctDNA in a 10-mL blood draw, mutant allele fraction (MAF; the percentage of ctDNA in comparison with the total cfDNA), and the likelihood of tumor detection using ctDNA genomic analysis (**Table 3**; ref. 13). We concluded that in general, these technologies are unlikely to detect tumors that are smaller than 10 mm

in diameter; the test will likely suffer a high rate of false negatives in early-stage cancer (13). The major reason behind the false negative results associated with ctDNA testing is sampling error, where the retrieval of sufficient ctDNA from the circulation to allow for robust data analysis and interpretation is a limiting factor.

Assay specificity

Similar to other cancer biomarkers, ctDNA testing may inevitably suffer from relatively low specificity in some cases. While investigators usually report sensitivities at 90%, 95%, or 98% specificity, these seemingly high thresholds may still not be enough when screening for less common cancers such as ovarian and pancreatic cancer. For example, if the cancer prevalence in the screened population is 1:400, a test with 100% sensitivity and 95% specificity will yield a PPV of 5%, meaning there is a 5% chance of a patient truly having cancer if the test is positive. This example highlights the well-known fact that screening tests must have extremely high specificity to ensure high PPV and avoid false positives. As we grow the genomic database of the healthy and cancer population, questions have been raised regarding just how cancer-specific are the genetic mutations found in ctDNA? Despite the expectation that cfDNA retrieved from normal individuals should not contain any genetic alterations, a plethora of new data shows otherwise (44). Mutations identified in normal tissues can be age related, may have functional significance in some tissues (such as the brain), and may be associated with precancerous lesions of unknown clinical significance (44). In addition to a questionable specificity to cancer, ctDNA analysis also usually cannot identify the affected organ/tissue and therefore cannot be used for diagnosis of specific cancer types (41). Cristiano and colleagues (41) reported a 61% chance of correct prediction of tumor localization, which is not far from tossing a coin. Therefore, a positive result by ctDNA analysis must be followed with a series of additional tests to corroborate the findings and identify the

Table 3. Approximate calculated relationship between mutant allele fraction and tumor characteristics.[a]

Mutant allele fraction[b], %	Tumor diameter, mm	Tumor weight (g) or volume (mL)
1	58	100
0.1	27	10
0.01	13	1
0.005	10	0.5
0.0025	8	0.25
0.0012	6	0.13
0.0006	5	0.06

[a]Data modified from ref. 13.
[b]MAF = ratio of ctDNA and cfDNA, expressed as a percentage.

tumor origin. This could be a serious issue, in addition to the cost of individualized genomic analysis and downstream testing, because the uncertainty of cancer specificity and tumor localization may induce patient anxiety and other side effects.

Recently, it has been brought to attention the impact of clonal hematopoiesis, a precancerous condition during which hematopoietic cells acquire mutations, has on complicating ctDNA analysis and compromising specificity. Ptashkin and colleagues found that 5% of all mutations found in tumors are due to clonal hematopoiesis and they drew attention that some patients may receive inappropriate therapies due to such errors (45). On the other hand, Razavi and colleagues found that more than 50% of identified mutations in ctDNA are due to clonal hematopoiesis, thus necessitating sequencing of white blood cell DNA to verify the original mutations (46, 47).

cfDNA fragmentation for cancer diagnosis

New work has shown that cfDNA fragmentation patterns can be used to diagnose cancer (14, 41). In this newest twist, it is hypothesized that the fragment lengths of cfDNA could be exploited for noninvasive genomic analysis and early detection of cancer (14, 41). Fragment length and position within the genome are both informative, as assessed by low-depth genomic sequencing of multiple 5 Mb regions, dispersed within the whole genome (for mechanisms of DNA fragmentation, please see refs. 14 and 41). The important observation is that ctDNA is, on average, slightly shorter, by about 3–6 bases, and cfDNA fragments derived from tumors are more variable in length compared with fragments in normal controls (14, 41).This revelation created a paradigm shift, from investigating one or a few genetic changes found only in ctDNA, to examining the whole spectrum of cfDNA fragment length variation and position to infer the presence of cancer (14, 41). By employing cfDNA fragment analysis alone, the reported sensitivities in 236 patients with breast, colorectal, lung, ovarian, pancreatic, gastric, or bile duct cancer ranged from 57% to 99% at 98% specificity (41). The sensitivity could be raised to 91% by combining this method with mutation-based ctDNA analysis (14, 41).

Calculating clinically relevant sensitivity, specificity, and lead time

Virtually every paper using cfDNA analysis for early cancer detection utilized clinically identified cancer cases and normal controls to calculate sensitivity and specificity. These approaches almost certainly overestimate the sensitivity of the test for early detection. Recent experimental data from the company GRAIL that wishes to develop a ctDNA test for cancer, can attest to these predictions and demonstrate the magnitude of the overestimation (48, 49). The GRAIL method was recently validated with clinically detected breast cancers and asymptomatic breast cancers detected by mammographic screening (49). These data have shown that the sensitivity drops by 4- to 5-fold, from 50% in clinically detected cancers to 10% in asymptomatic, mammographically-detected cancers (49).

GRAIL recently presented their latest validation data with an independent patient set at the inaugural American Society of Clinical Oncology (ASCO) Breakthrough conference in the fall of 2019 (50). It is important to mention that the patients in this set have also already been diagnosed clinically, so the experiment does not closely mimic a screening scenario which usually includes asymptomatic individuals. In this new validation set, the diagnostic specificity was raised to 99% (from 95%), a significant achievement. The overall sensitivity for multiple cancer sites was 55%, but for early-stage tumors (stage I) that are least likely to be detected clinically or through screening, the average detection rate (sensitivity) was only 18%.

Another way to validate the sensitivity of new technologies, such as cfDNA, for early cancer detection would be to analyze samples collected longitudinally, as part of prospective randomized trials such as the Prostate, Lung, Colorectal, and Ovarian (PLCO) Cancer Screening Trial (51). This large trial sponsored by the NCI is an example of an ideal sample set for determining the effects of screening on cancer-related mortality (primary end point) and other secondary endpoints (i.e., progression-free survival, etc.), in men and women ages 55 to 74 (51).

Under the auspices of EDRN, we used samples from the PLCO (180 ovarian cancer cases and 660 controls) to evaluate and rank 49 ovarian cancer biomarkers for early, preclinical ovarian cancer diagnosis (52). We found that none of these biomarkers were effective, alone or in combinations, in detecting asymptomatic disease (52). The conventional ovarian cancer biomarker CA125, which has about 80%–90% sensitivity in detecting clinical disease, exhibited less than 50% sensitivity in asymptomatic patients (52). This finding underlines the challenges of using biomarkers to detect cancer at an early stage, where individuals are usually asymptomatic. The lead time from asymptomatic to clinical diagnosis must also be long enough (e.g., more than 6 months), if earlier therapeutic interventions are expected to have some positive effect in overall and disease-free survival. Shorter lead times for certain cancers would render early detection useless, and even high-performing biomarkers for these cancers may be of low clinical value if there are no significant benefits in patient survival associated with early detection.

Reporting the mutant allele fraction of specimens

We previously correlated MAF (the abundance of mutant alleles associated with ctDNA in comparison with normal alleles found in the total cfDNA in a plasma sample) with other tumor characteristics such as size and weight (**Table 3**; ref. 13). We concluded that when there is less than one copy of ctDNA intermixed with 10,000 copies of normal cfDNA (MAF of 0.01%), tumor diagnosis is unlikely due to sampling error; in such case, there will be no retrieval of even a single copy of ctDNA in a 10-mL blood draw (13). At a MAF of 0.01%, we calculated that the tumor will be about 12 mm in diameter (13). In most published studies, the MAF is 0.1% or higher, which yields favorable assay sensitivity (39). However, the tumor diameter will likely be more than 27 mm with a MAF of 0.1%, and the tumor will be easily detectable by imaging. In the latest report by Cristiano and colleagues, the MAF of the cancers studied is not mentioned (41). Some studies stratify patients according to stage, but stage and tumor size do not always closely correlate, because some relatively large tumors remain localized and could be classified as early stage. We strongly recommend that in future studies, the MAF of all utilized samples is mentioned to provide an idea of tumor size, to assess the utility of the biomarker compared with imaging tests.

Artificial intelligence algorithms

As research makes strong inroads into cancer detection, prognosis, and monitoring with the new field of ctDNA testing, the amount of sequencing data we accumulate grows exponentially, which allows for sophisticated analysis using artificial intelligence (AI) and machine learning. Studies have begun to use AI algorithms to combine not only information on MAF, but also cfDNA fragment length and position, to provide a holistic prediction of the presence of cancer (14, 39). In theory, cfDNA-based tests could work even if there is less than one representative complete genome in a 10-mL blood draw because the information is derived from a fraction of genomic DNA regions

Table 4. Some scientific debates and their outcomes.

Technology/theme	Principal investigators/ company	Date	Journal (reference)	Opposition reference	Outcome of original article/work
Ovarian cancer diagnostics using proteomic patterns	L. Liotta/E. Petricoin	2006	Lancet (57)	(58)	Independent validation by EDRN failed but paper was not retracted
Serum peptidomics for cancer diagnostics	P. Tempst	2006	Journal of Clinical Investigation (59)	(60)	No follow-up
Cancer diagnostics using nuclear matrix proteins	R. Getzenberg	2002–2010	Cancer Research (61)	(62)	Many relevant papers were retracted
Direct to consumer point-of-care diagnostics	Theranos	2015	No publications	(63)	Company shut down; ongoing criminal investigations
Cancer diagnostics using tumor-educated platelets	M.G. Best	2015–2016	Cancer cell (64)	(65)	Uncertain
ctDNA as an early cancer diagnostic test	V. Velculescu/ N. Papadopoulos (GRAIL, THRIVE, and others)	2017–2019	Science Translational Medicine, Science, Nature (39–42)	(13, 43)	Data gathering for commercialization
Wellness testing-P4 Medicine	L. Hood/Arivale	2017–2018	Nature Biotechnology (66)	(55–57, 67)	Arivale filed for bankruptcy (2019)
Role of microbiome in immunotherapy	V. Gopalakrishnan/ V. Matson/B. Routy	2019	Science (68–70)	(71, 72)	Uncertain

Abbreviation: EDRN, Early Detection Research Network (NCI/NIH Initiative).

originating from normal and cancer cells (41). In general, this method covers approximately 2.6 GB of DNA (504 windows of 5 Mb each; approximately 20,000 sequencing reads per window; ref. 41). In total, the sequencing generates about 10,000,000 reads at a depth of about $1\times$. When the mutant allele fraction is 1:10,000 (0.01%), approximately 10,000,000 reads will represent normal DNA and only 1,000 reads will represent cancer-derived DNA. Although we do not know the details of the algorithm, one would expect that as the MAF decreases, the representation of the ctDNA will also decrease, along with the strength of the AI prediction algorithm. Consequently, it may be necessary to demonstrate the strength of the prediction algorithm at various MAF levels to avoid overestimation of ctDNA sensitivity and to understand the variability in the data. Although experimental data is lacking, we predict that at a MAF of 0.001%, when tumors are expected to be approximately 6 mm in diameter, the mutant DNA reads will be too few for AI algorithms to make an accurate prediction. All in all, coupled with sampling error, as we mentioned above, where ctDNA may not be retrieved at sufficient amounts in a 10-mL blood draw, the current uncertainty for AI programs to perform effectively at extremely low MAF (<0.01%) constitutes another roadblock of ctDNA testing for early cancer detection.

Challenging False Discovery and the Limitations in Biomarker Research

Is it a good idea to challenge questionable data? Science progresses faster if published results are debated in the appropriate fora. In most cases, differences in opinions originate from honest oversights that could produce false discoveries (data fabrication is a rare phenomenon). However, highlighting mistakes and flaws in published literature, especially reports on new and exciting biomarkers that seem to hold promise for revolutionizing cancer care, takes courage from the challengers, who put their reputation on the line. In short, there is no room for mistakes when criticizing other people's mistakes. For this reason, challenges of published data are not very common. On the other hand, those who published problematic data are very reluctant to accept their mistakes and retract their papers, due to the stigma and possible career consequences. Naturally, a large number of false reports still resides in the scientific literature.

In biomarker research, false discovery can be concealed by various explanations such as different assays, different algorithms, different samples, etc. At the end of the day, the ultimate judge is time. False discoveries never reach the clinic or can lead to start-ups that eventually fail, sadly wasting funds and energy. We provided many examples of presumed false discoveries in previous papers (**Table 4**; refs. 5, 15). Even though many of these "discoveries" have been disproved, the original papers have not been retracted and some are still highly cited.

One of our own proposals to reduce false discovery is the 5- or 10-year reflection, which mandates that authors of significant papers in biomarkers and other fields write a 5- to 10-year perspective on their initial discoveries (20, 21). While authors are very keen to speak to the media at the time of publication, they are often unwilling to examine the fate of their discoveries 5–10 years later, for obvious reasons. There is room for meta-analyses which can track the fate of seemingly revolutionary discoveries (including biomarkers) to identify the reasons for their ultimate failure and the lessons learned.

Conclusions and Future Perspectives

The road to new cancer biomarker discovery and validation is plagued with barriers that are hard to cross. Regardless of the newest and most advanced technology that is employed, judging the true value of a new biomarker returns to the familiar factors of clinical importance (unmet needs), sensitivity, specificity, PPV, and NPV for the cancer type at hand. The extensive interrogation of the cancer genome and tumor heterogeneity has opened a whole new realm for cancer biomarker research, where personalized biomarkers, in the form of a repository of protein or gene-based markers that can be used to screen each patient, may be the future.

Among these possibilities lies the new concept of ctDNA testing, where its potential applications as a groundbreaking liquid biopsy assay spans from cancer detection, assessing therapeutic response, to

monitoring tumor growth and relapse. Equipped with a sharpened understanding of the utility and drawbacks of cfDNA as a new cancer biomarker, it is undoubtable that this field holds capacity for entering routine clinical care in the future if the current limitations are addressed appropriately with careful study design, larger datasets, and continuous technological progression in sequencing and bioinformatics methods. Regarding ctDNA, we suggest that the sensitivity should be determined using samples with a known MAF, with a focus on early-stage cancers, if possible, to examine its feasibility for early detection. In general, the specificity of new biomarkers should be assessed by including age- and sex-matched patients with benign and inflammatory diseases, and sample size calculations should take into account the prevalence of the cancer type.

Additional steps for validating new biomarkers would be to test prediagnostic samples available from prospective clinical trials like the PLCO, to calculate both sensitivity and lead time before clinical diagnosis. Furthermore, the lead time achieved with early detection will tell us whether patient outcomes can be influenced through earlier therapies. Although these stepwise approaches are intuitive to biomarker discovery, screening experiences, such as with ovarian cancer, have not supported this expectation (53). Another important issue is to make sure that screening, particularly in low prevalence cancers, avoids overdiagnosis and overtreatment, as outlined by us and others elsewhere (54–57). The best way to consolidate the success rate of new biomarker discoveries is to encourage respectful scientific debates surrounding new breakthroughs. The impediments encountered in biomarker development are less likely to hamper new discoveries if they are taken into account and implemented in the study design along each step of the pipeline. In the end, the recipe for fulfilling the mandate of the EDRN is an intricate balance of challenging common oversights and voicing concerns over flaws in research, which will lead to decreased incidences of false discovery as well as better usage of funding and time, by focusing resources on high-performance biomarkers that will eventually transform routine cancer patient care.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

References

1. Etzioni R, Urban N, Ramsey S, McIntosh M, Schwartz S, Reid B, et al. The case for early detection. Nat Rev Cancer 2003;3:243–52.
2. Srivastava S, Rossi SC. Early detection research program at the NCI. Int J Cancer 1996;69:35–7.
3. Kulasingam V, Diamandis EP. Strategies for discovering novel cancer biomarkers through utilization of emerging technologies. Nat Clin Pract Oncol 2008;5: 588–99.
4. Kulasingam V, Pavlou MP, Diamandis EP. Integrating high-throughput technologies in the quest for effective biomarkers for ovarian cancer. Nat Rev Cancer 2010;10:371–8.
5. Diamandis EP. Cancer biomarkers: can we turn recent failures into success? J Natl Cancer Inst 2010;102:1462–7.
6. Diamandis EP. Towards identification of true cancer biomarkers. BMC Med 2014;12:156.
7. Diamandis EP. Analysis of serum proteomic patterns for early cancer diagnosis: drawing attention to potential problems. J Natl Cancer Inst 2004;96:353–6.
8. Makawita S, Diamandis EP. The bottleneck in the cancer biomarker pipeline and protein quantification through mass spectrometry-based approaches: current strategies for candidate verification. Clin Chem 2010;56:212–22.
9. Rittenhouse H, Blase A, Shamel B, Schalken J, Groskopf J. The long and winding road to FDA approval of a novel prostate cancer test: our story. Clin Chem 2013; 59:32–4.
10. Ioannidis JPA, Bossuyt PMM. Waste, leaks, and failures in the biomarker pipeline. Clin Chem 2017;63:963–72.
11. Ioannidis JPA. Biomarker failures. Clin Chem 2013;59:202–4.
12. Bettegowda C, Sausen M, Leary RJ, Kinde I, Wang Y, Agrawal N, et al. Detection of circulating tumor DNA in early- and late-stage human malignancies. Sci Transl Med 2014;6:224ra24.
13. Fiala C, Diamandis EP. Utility of circulating tumor DNA in cancer diagnostics with emphasis on early detection. BMC Med 2018;16:166.
14. Mouliere F, Chandrananda D, Piskorz AM, Moore EK, Morris J, Ahlborn LB, et al. Enhanced detection of circulating tumor DNA by fragment size analysis. Sci Transl Med 2018;10. pii: eaat4921.
15. Diamandis EP. The failure of protein cancer biomarkers to reach the clinic: why, and what can be done to address the problem? BMC Med 2012;10:87.
16. Schully SD, Carrick DM, Mechanic LE, Srivastava S, Anderson GL, Baron JA, et al. Leveraging biospecimen resources for discovery or validation of markers for early cancer detection. J Natl Cancer Inst 2015;107. pii: djv012.
17. Ransohoff DF. Bias as a threat to the validity of cancer molecular-marker research. Nat Rev Cancer 2005;5:142–9.
18. Diamandis EP, Fiala C. Make researchers revisit past publications to improve reproducibility. F1000Res 2017;6:1717.
19. Wallach JD, Boyack KW, Ioannidis JPA. Reproducible research practices, transparency, and open access data in the biomedical literature, 2015–2017. PLoS Biol 2018;16:e2006930.
20. Fiala C, Diamandis EP. Benign and malignant scientific irreproducibility. Clin Biochem 2018;55:1–2.
21. Fiala C, Diamandis EP. How to reduce scientific irreproducibility: the 5-year reflection. Clin Chem Lab 2017;55:1845–8.
22. Dagogo-Jack I, Shaw AT. Tumour heterogeneity and resistance to cancer therapies. Nat Rev Clin Oncol 2018;15:81–94.
23. Yuan Y. Spatial heterogeneity in the tumor microenvironment. Cold Spring Harb Perspect Med 2016;6. pii: a026583.
24. Marusyk A, Polyak K. Tumor heterogeneity: causes and consequences. Biochim Biophys Acta 2010;1805:105–17.
25. Lawson DA, Kessenbrock K, Davis RT, Pervolarakis N, Werb Z. Tumour heterogeneity and metastasis at single-cell resolution. Nat Cell Biol 2018;20: 1349–60.
26. Prasetyanti PR, Medema JP. Intra-tumor heterogeneity from a cancer stem cell perspective. Mol Cancer 2017;16:41.
27. McGranahan N, Swanton C. Clonal heterogeneity and tumor evolution: past, present, and the future. Cell 2017;168:613–28.
28. Appierto V, Di Cosimo S, Reduzzi C, Pala V, Cappelletti V, Daidone MG. How to study and overcome tumor heterogeneity with circulating biomarkers: the breast cancer case. Semin Cancer Biol 2017;44:106–16.
29. Gerlinger M, Rowan AJ, Horswell S, Math M, Larkin J, Endesfelder D, et al. Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. N Engl J Med 2012;366:883–92.
30. Jamal-Hanjani M, Quezada SA, Larkin J, Swanton C. Translational implications of tumor heterogeneity. Clin Cancer Res 2015;21:1258–66.
31. Gu G, Dustin D, Fuqua SA. Targeted therapy for breast cancer and molecular mechanisms of resistance to treatment. Curr Opin Pharmacol 2016;31:97–103.
32. Pishvaian MJ, Brody JR. Therapeutic implications of molecular subtyping for pancreatic cancer. Oncology 2017;31:159–66, 168.
33. Yu Y. Molecular classification and precision therapy of cancer: immune checkpoint inhibitors. Front Med 2018;12:229–35.
34. Sinicrope FA, Okamoto K, Kasi PM, Kawakami H. Molecular biomarkers in the personalized treatment of colorectal cancer. Clin Gastroenterol Hepatol 2016;14: 651–8.
35. Kulasingam V, Prassas I, Diamandis EP. Towards personalized tumor markers. NPJ Precis Oncol 2017;1:17.
36. Wan JCM, Massie C, Garcia-Corbacho J, Mouliere F, Brenton JD, Caldas C, et al. Liquid biopsies come of age: towards implementation of circulating tumour DNA. Nat Rev Cancer 2017;17:223–38.
37. Corcoran RB, Chabner BA. Application of cell-free DNA analysis to cancer treatment. N Engl J Med 2018;379:1754–65.

38. Aravanis AM, Lee M, Klausner RD. Next-generation sequencing of circulating tumor DNA for early cancer detection. Cell 2017;168:571–4.
39. Phallen J, Sausen M, Adleff V, Leal A, Hruban C, White J, et al. Direct detection of early-stage cancers using circulating tumor DNA. Sci Transl Med 2017;16. pii: eaan2415.
40. Cohen JD, Li L, Wang Y, Thoburn C, Afsari B, Danilova L, et al. Detection and localization of surgically resectable cancers with a multi-analyte blood test. Science 2018;359:926–30.
41. Cristiano S, Leal A, Phallen J, Fiksel J, Adleff V, Bruhm DC, et al. Genome-wide cell-free DNA fragmentation in patients with cancer. Nature 2019;570: 385–9.
42. Shen SY, Singhania R, Fehringer G, Chakravarthy A, Roehrl MHA, Chadwick D, et al. Sensitive tumour detection and classification using plasma cell-free DNA methylomes. Nature 2018;563:579–83.
43. Fiala C, Kulasingam V, Diamandis EP. Circulating tumor DNA for early cancer detection. J Appl Lab Med 2018;2:300–13.
44. Tomasetti C. Mutated clones are the new normal. Science 2019;364:938–9.
45. Ptashkin RN, Mandelker DL, Coombs CC, Bolton K, Yelskaya Z, Hyman DM, et al. Prevalence of clonal hematopoiesis mutations in tumor-only clinical genomic profiling of solid tumors. JAMA Oncol 2018;4:1589–93.
46. Razavi P, Li BT, Brown DN, Jung B, Hubbell E, Shen R, et al. High-intensity sequencing reveals the sources of plasma circulating cell-free DNA variants. Nat Med 2019;25:1928–37.
47. Fiala C, Diamandis EP. Circulating tumor DNA (ctDNA) is not a good proxy for liquid biopsies of tumor tissues for early detection. Clin Chem Lab Med 2020 Mar 11 [Epub ahead of print].
48. Klein EA, Hubbell E, Maddala T, Aravanis A, Beausang JF, Filippova D. Development of a comprehensive cell-free DNA (cfDNA) assay for early detection of multiple tumor types: the Circulating Cell-free Genome Atlas (CCGA) study. J Clin Oncol 36, 2018 (suppl; abstr 12021).
49. Liu MC, Maddala T, Aravanis A, Hubbell E, Beausang JF, Filippova D, et al. Breast cancer cell-free DNA (cfDNA) profiles reflect underlying tumor biology: the Circulating Cell-Free Genome Atlas (CCGA) study. J Clin Oncol 36, 2018 (suppl; abstr 536).
50. Oxnard GR, Klein EA, Seiden M, Hubbell E, Venn O, Jamshidi A, et al. Simultaneous multi-cancer detection and tissue of origin (TOO) localization using targeted bisulfite sequencing of plasma cell-free DNA (cfDNA). J Glob Oncol 2019;5:44.
51. Gohagan JK, Prorok PC, Greenwald P, Kramer BS. The PLCO cancer screening trial: background, goals, organization, operations, results. Rev Recent Clin Trials 2015;10:173–80.
52. Cramer DW, Bast RC, Berg CD, Diamandis EP, Godwin AK, Hartge P, et al. Ovarian cancer biomarker performance in prostate, lung, colorectal, and ovarian cancer screening trial specimens. Cancer Prev Res 2011;4:365–74.
53. Jacobs IJ, Menon U, Ryan A, Gentry-Maharaj A, Burnell M, Kalsi JK, et al. Ovarian cancer screening and mortality in the UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS): a randomised controlled trial. Lancet 2016;387: 945–56.
54. Diamandis EP, Li M. The side effects of translational omics: overtesting, overdiagnosis, overtreatment. Clin Chem Lab Med 2016;54:389–96.
55. Srivastava S, Koay EJ, Borowsky AD, De Marzo AM, Ghosh S, Wagner PD, et al. Cancer overdiagnosis: a biological challenge and clinical dilemma. Nat Rev Cancer 2019;19:349–58.
56. Fiala C, Taher J, Diamandis E. P4 medicine or O4 medicine? Hippocrates provides the answer. J Appl Lab Med 2018;3:028613.
57. Petricoin EF, Ardekani AM, Hitt BA, Levine PJ, Fusaro VA, Steinberg SM, et al. Use of proteomic patterns in serum to identify ovarian cancer. Lancet 2002;359: 572–7.
58. Diamandis EP. Point: proteomic patterns in biological fluids: do they represent the future of cancer diagnostics? Clin Chem 2003;49:1272–5.
59. Villanueva J, Shaffer DR, Philip J, Chaparro CA, Erdjument-Bromage H, Olshen AB, et al. Differential exoprotease activities confer tumor-specific serum peptidome patterns. J Clin Invest 2006;116:271–84.
60. Diamandis EP. Letter to the Editor about Differential exoprotease activities confer tumor-specific serum peptidome. J Clin Invest 2006;116:1.
61. Brünagel G, Vietmeier BN, Bauer AJ, Schoen RE, Getzenberg RH. Identification of nuclear matrix protein alterations associated with human colon cancer. Cancer Res 2002;62:2437–42.
62. Diamandis EP. Early prostate cancer antigen-2: a controversial prostate cancer biomarker? Clin Chem 2010;56:542–4.
63. Diamandis EP. Theranos phenomenon: promises and fallacies. Clin Chem Lab Med 2015;53:989–93.
64. Best MG, Sol N, Kooi I, Tannous J, Westerman BA, Rustenburg F, et al. RNA-seq of tumor-educated platelets enables blood-based pan-cancer, multiclass, and molecular pathway cancer diagnostics. Cancer Cell 2015;28:666–76.
65. Diamandis EP. A word of caution on new and revolutionary diagnostic tests. Cancer Cell 2016;29:141–2.
66. Price ND, Magis AT, Earls JC, Glusman G, Levy R, Lausted C, et al. A wellness study of 108 individuals using personal, dense, dynamic data clouds. Nat Biotechnol 2017;35:747–56.
67. Fiala C, Taher J, Diamandis EP. Benefits and harms of wellness initiatives. Clin Chem Lab Med 2019;57:1494–500.
68. Gopalakrishnan V, Spencer CN, Nezi L, Reuben A, Andrews MC, Karpinets TV, et al. Gut microbiome modulates response to anti-PD-1 immunotherapy in melanoma patients. Science 2018;359:97–103.
69. Matson V, Fessler J, Bao R, Chongsuwat T, Zha Y, Alegre M-L, et al. The commensal microbiome is associated with anti-PD-1 efficacy in metastatic melanoma patients. Science 2018;359:104–8.
70. Routy B, Le Chatelier E, Derosa L, Duong CPM, Alou MT, Daillère R, et al. Gut microbiome influences efficacy of PD-1-based immunotherapy against epithelial tumors. Science 2018;359:91–7.
71. Janket S-J, Ackerson LK, Diamandis E. Simpson's paradox in proof-of-concept studies. Nat Med 2019;25:1640.
72. Janket S-J, Ackerson LK, Diamandis EP. Gut microbiotas and immune checkpoint inhibitor therapy response: a causal or coincidental relationship? Clin Chem Lab Med 2019;58:18–24.