Published in final edited form as:

Ann Biomed Eng. 2020 October ; 48(10): 2377–2399. doi:10.1007/s10439-020-02593-y.

Scalable signature-based molecular diagnostics through on-chip biomarker profiling coupled with machine learning

John Molinski¹, Amogha Tadimety¹, Alison Burklund¹, John X.J. Zhang^{1,2,*}

¹:Thayer School of Engineering at Dartmouth, Hanover, NH

²:Norris Cotton Cancer Center, Dartmouth Hitchcock Medical Center, Lebanon, NH

Abstract

Molecular diagnostics have traditionally relied on discrete biological substances as diagnostic markers. In recent years however, advances in on-chip biomarker screening technologies and data analytics have enabled signature-based diagnostics. Such diagnostics aim to utilize unique combinations of multiple biomarkers or diagnostic ‘fingerprints’ rather than discrete analyte measurements. This approach has shown to improve both diagnostic accuracy and diagnostic specificity. In this review, signature-based diagnostics enabled by microfluidic and micro-/nano-technologies will be reviewed with a focus on device design and data analysis pipelines and methodologies. With increasing amounts of data available from microfluidic biomarker screening, isolation, and detection platforms, advanced data handling and analytics approaches can be employed. Thus, current data analysis approaches including machine learning and recent advances with image processing, along with potential future directions will be explored. Lastly, the needs and gaps in current literature will be elucidated to inform future efforts towards development of molecular diagnostics and biomarker screening technologies.

Keywords

Molecular profiling; Biomarker screening; Micro-/Nano- technologies; Advanced data analytics

1. Introduction:

Molecular diagnostics are critical tools for clinical decision-making in the treatment of both infectious and chronic disease¹. The choice of disease-specific biomarkers has long been fundamental to the design and development of diagnostic devices and assays. In contrast, signature-based diagnostics have emerged as an attractive alternative to those based on traditional biomarkers. Signature-based diagnostics utilize biomarker signatures or multiple markers in combination, effectively creating a ‘diagnostic fingerprint’ for a disease state of

Terms of use and reuse: academic research for non-commercial purposes, see here for full terms. <https://www.springer.com/aam-terms-v1>

*Corresponding author. **Mailing address:** Thayer School of Engineering at Dartmouth, 14 Engineering Drive, Hanover, NH, 03755, USA. **Phone:** 1-603- 646-8787. john.zhang@dartmouth.edu.

Publisher's Disclaimer: This Author Accepted Manuscript is a PDF file of an unedited peer-reviewed manuscript that has been accepted for publication but has not been copyedited or corrected. The official version of record that is published in the journal is kept up to date and so may therefore differ from this version.

interest. Patient-specific data resulting from a disease signature enables targeted and personalized medicine, streamlines the diagnostic pipeline, and has the potential to dramatically improve patient outcomes².

Biomarker signatures are a panel of distinct yet often interrelated biomarkers (typically three or more) which represent a disease state of interest. Relationships between biomarkers within a signature can be simple (i.e. concentration of each marker) or complex (i.e. a relative expression of each marker with respect to one another). An example of the former arises when individual biomarkers are selected and an aggregate sum of their concentration or expression is used as a diagnostic metric. Contrarily, an example of the latter is when heatmaps of biomarker expression are presented as a singular marker. At present, biomarker signatures are most often made up of cellular and vesicular proteins^{3–5} or nucleic acids such as microRNA^{6–9} and mRNAs¹⁰. Signature-based diagnostics have been predominately investigated for the diagnosis and monitoring of cancer^{5,10–24}, but have also been utilized for the diagnosis of Alzheimer's^{25–31}, traumatic brain injury^{9,32}, and coronary artery disease^{4,33}. Though promising preliminary results have been demonstrated, diagnostic signatures inherently increase diagnostic complexity and clinical validation in large cohorts are required to assess their clinical viability.

The development of signature-based diagnostic platforms requires sophisticated device designs and downstream data analysis techniques that are able to output these potentially clinically significant signature profiles. Microfluidic and micro-/nano- technologies are particularly suited for the specific manipulation and isolation of biological molecules^{34–36}, with characteristic dimensions matching that of the biomolecules. Advances in device design and detection mechanisms have allowed for highly parallelized analysis, which has increased the feasibility of on-chip molecular profiling and signature-based diagnostics. When coupled with data processing and analysis tools, these platforms have been demonstrated to accurately classify disease states^{5,9} from healthy patients and enable and potentially enable highly data-driven clinical decision-making¹⁰. As such, machine learning and image processing techniques have a newfound role as fundamental components of diagnostic systems given their applicability to rapidly and accurately analyze these complex datasets³⁷. By leveraging these advanced data processing techniques, signature-based diagnostics have the potential to make dramatic contributions to personalized medicine, thus improving patient outcomes³⁸. With these advanced data processing techniques however, effort should be made to increase transparency in data processing workflows and machine learning models to ensure translatability for different devices and potentially diseases.

This field of coupling molecular profiling and advanced analysis tools is still relatively new, and there is significant opportunity for continued research. Currently, there are no widely accepted device designs nor downstream data processing methodologies. These tools are urgently needed to avoid ambiguity in the decision-making process, and to provide a universal comparison metric for diagnostic development. Within diagnostic development, aside from device performance, advances which enable scalability of device production must be considered to enable clinical translation. This review aims to highlight the recent progress in both device design and data analytics to allow for fully integrated, signature-based molecular diagnostic devices. The paper will start with an overview of device and data

analytics methods, and then will review a selection of microfluidic and micro-/nano-technologies that integrate advanced analysis.

1.1 Biomarker selection in signature diagnostics:

The diagnostic accuracy of a biomarker signature relies both on the individual markers chosen and the interrelationship of these markers, making biomarker selection a critically important decision. Oftentimes the biomarker selection is based on differences in concentration^{39,40} and/or relative expression in healthy versus diseased patients^{41–43}. In contrast to singular marker diagnostics the advent of signature-based biomarker panels can provide a unique and multidimensional viewpoint on the disease state of interest with the potential to probe complex diseases such as cancer. At present, such panels are still early in the discovery phase and evaluation of panels are focused primarily on diagnostic accuracy. If to be used as a clinically meaningful metric, selected signatures must undergo rigorous clinical validation to examine potential pitfalls and benchmarked to singular marker diagnostics using standard evaluation metrics.

Enveloped molecules, such as cells and vesicles, are a common target as they provide both extra- and intra-cellular/vesicular markers for analysis. Extra-cellular/vesicular markers within enveloped molecules are often surface proteins, which can be targeted directly via antibody-antigen interactions for isolation and/or detection. Intra-cellular/vesicular markers include genetic cargo, such as miRNA and mRNA^{19,42,44–47}, and intra-cellular/vesicular proteins. The combination of capture using extra-cellular/vesicular markers with downstream molecular profiling of enveloped molecules aids the development of unique biomarker signatures for specific diseases and disease states.

Depending on the enveloped biomarker molecule chosen, varying levels of purification and preprocessing are required prior to detection and analysis. For example, some commonly exploited enveloped biomarkers for diagnosis and characterization of cancer include circulating tumor cells (CTCs) and exosomes. In the case of circulating tumor cells (CTCs), their low concentration presents a technical challenge, with typically only 1–10 CTCs per mL of whole blood⁴⁴. Moreover, because white blood cells are similar in size, it is difficult to isolate CTCs based on size alone. Thus, capture methods generally target overexpressed surface proteins specific to CTCs such as EpCAM⁴⁵, CD44⁴⁶, and CD81⁴⁷. In spite of these challenges, CTCs are widely studied biomarkers for cancer diagnostics due to their abundance of molecular cargo.

Similar to CTCs, tumor-derived exosomes are also enveloped markers and carry a variety of host-cell-specific molecular cargo, such as microRNA, proteins, mRNA and DNA^{48–51}. Exosomes are nanoscale sized phospholipid extracellular vesicles, typically 30–150 nm in diameter. These vesicles were once thought to be cellular waste⁵², but recent work has elucidated their importance in mediating cell-to-cell signaling interactions^{21,49,52,53}, along with their potential link to tumor progression and metastasis^{54,55}. Prior work has elucidated many surface proteins which are overexpressed in tumor-derived subpopulations^{48,56}, such as CD24⁵⁷, survivin⁵⁸, and EpCAM⁵⁹. As in the case of CTCs, the specific overexpressed surface proteins can vary with cancer subtype being targeted. Aside from surface proteins, previous groups have shown that intravesicular markers such as miRNA and mRNAs are

also over/under-expressed in cancer subtypes, providing a robust suite of makers available for selection within enveloped molecules⁶⁰. Exosome capture and analysis has presented significant challenges over CTC's due to their size. Though modified techniques previously used for CTC's have been utilized often entirely new approaches are introduced.

1.2 Microfluidics for biomarker screening and detection:

Microfluidic devices allow for manipulation and control of minute volumes of fluids through use of small channels on the micron size scale. Typically, soft polymers such as polydimethylsiloxane (PDMS) are used in fabrication due to the versatility and potential for mass fabrication. Microfluidic devices have long been utilized for biomarker isolation due to the ability to design devices at the same scale of various target molecules^{61,62}. The use of microfluidics within the clinical setting has also been on the rise due to the promise of controlled biofluid handling with small (<μL) volumes and a clinically relevant throughput. More recent advances in fabrication and design have allowed for highly parallelized on-chip operation amenable to multiplexing and the development of signature-based diagnostic platforms. In these devices, often each channel or chamber is made to select a specific biomarker of interest^{63–65}, allowing for the isolation of multiple markers simultaneously. Although this approach to multiplexed diagnostics is achievable, added fabrication concerns can quickly challenge clinical translation and implementation. Moreover, due to the added complexity in data output, advanced data processing techniques such as image processing or machine learning^{20,66–68} often need to be incorporated with the device to enable accurate and timely analysis. Thus, scalable device design must be a focus in the development of signature-based diagnostic devices to enable clinical translation and address current limitations.

Clinical microfluidic systems aim to both isolate and detect biomarkers of interest to enable downstream analysis and quantification, ultimately to inform a diagnosis^{69–71}. Biomarker enrichment platforms aim to differentiate and separate biomolecules of interest from complex biofluids that are commonly encountered within a clinical setting. Isolation can involve targeted enrichment of the biomolecule (positive enrichment) or targeted depletion of interfering molecules (negative enrichment). In devices aimed at discrete biomolecule isolation, a host of modalities have been explored, including dielectrophoretic^{72,73}, acoustic^{74–77}, geometric^{78–82}, and immunomagnetic/immunoaffinity^{12,83–88} approaches. Broadly, these separation methods can be characterized into either chemical and/or physical isolation methods. Chemical enrichment methods target distinct biochemical properties such as molecular heterogeneity and overexpressed surface markers including proteins. Chemical isolation often utilizes immunoaffinity or aptamer-based separation methods to target these unique biochemical moieties. This separation modality is highly specific but often involves complex device chemistry and is relatively high cost. Alternatively, physical separation methods utilize differences such as size, deformability and compressibility within biomolecules of interest. Isolation can be achieved by the introduction of an external force such as an acoustic force or electric field, or through optimization of device geometry to capitalize on fluid inertial forces. Both chemical and physical approaches complement microfluidic workflows and allow for discrete biomarker isolation.

For use in signature-based and highly parallelized devices, many of these aforementioned approaches are not scalable without considerable alterations in device design or fabrication. As such, microfluidic devices developed for highly parallelized analysis often leverage immunoaffinity separation, due to the ability to target multiple markers on enveloped biomolecules, such as circulating tumor cells and extracellular vesicles. Moreover, with this approach, isolation of target biomolecules can be highly selective through careful choice and targeting of key recognition molecules. Further, integration of immunoaffinity approaches with additional separation modalities can increase capture efficiency and lower limits of detection. To achieve this, optimized device geometries have been combined with immunoaffinity capture to optimize substrate interaction and thus, capture efficiency. These methods for biomarker isolation provide are a necessary step to allow for downstream detection, quantification, and characterization.

Another important component of microfluidic diagnostic systems is the ability to detect and quantify the isolated analyte(s). Detection of target markers has been explored by means of fluorescence⁵, plasmonic/photonic materials^{89,90}, or through electric/electrochemical means^{91,92}. In the case of signature-based diagnostics, multiple markers are targeted so that multiple quantifiable outputs are generated. Here, concentration of each marker, or relative expression of the panel of markers can provide a biomarker signature that reflects a specific disease state^{93–95}. Early microfluidic systems focused solely on biomarker screening, utilizing off-chip downstream –omic profiling for analysis^{96,97}. More recent advances have combined on-chip screening with on-chip detection, allowing for fully integrated analytical systems, reducing manual intervention and time-to-result^{70,98–100}. Still however, following isolation and detection, data processing integration is vital to extract meaningful results from the resulting data. This integration of isolation, detection, and data analysis has the potential to create clinically useful, micro-scale, signature-based diagnostic platforms.

2. Advanced data analytics towards signature-based diagnostics:

Advanced data processing and analysis tools are capable of parsing though large and complex biomarker datasets, making these tools especially relevant to signature-based diagnostic platforms. The following sections will highlight two data analysis methods that have gained considerable interest in this field: machine learning and image processing.

2.1 Signature-based diagnostics enabled by machine learning:

Machine learning is a subset of artificial intelligence involving algorithms which can ‘learn’ from data by extracting patterns and information without explicit programming. These algorithms have become increasingly popular due to their unparalleled ability to provide insight into extremely complex datasets. ‘Learning’ within machine learning generally falls under one of two categories, namely supervised or unsupervised learning. In the former, training data is labeled, and the algorithm attempts to optimize the model such that classification error is minimized, thus supervised learning is ideally suited for classification problems. Conversely, in unsupervised algorithms training data is unlabeled and the algorithm attempts to ‘learn’ patterns within the data via feature extraction, which it can then use to cluster or associate like data. A limitation in unsupervised learning models is the

‘black box’ nature of these models where the features the model attends to are unknown. In a clinical setting such features can provide potential for biomarker discovery and diagnostic value however effort must be made to elucidate model workings to ensure underlying scientific validation before clinical use. Commonly explored machine learning algorithms explored include neural and deep networks, support vector machines, K-means clustering, linear discriminant analysis (LDA), random forest classifiers and decision trees⁶⁶. The next sections will cover the most commonly used machine learning techniques within signature-based diagnostic technologies and highlight recent advances in deep learning poised to make future contributions.

2.1.1 Dimensionality reduction—Linear discriminant analysis (LDA) is a dimensionality reduction technique used within multi-class classification problems which aims to separate data within different classes based on a projection of the data to a linear axis. Within LDA, for a projection to be made two criteria must be met. Namely, the distance between the means of two classes must be maximized within the, while the variance within each class is minimized¹⁰¹. Depending on the number of classes, the result of this technique is a set of linear discriminators, or a set of functions which represent the axes of maximum separability between each class. Predictions made from this model are done so in a probabilistic manner. When new data is introduced, its classification is determined by examining the probability that that data belongs to each class and selecting the class of highest probability. Many groups have begun to incorporate LDA to build classifiers through which disease classifications can be made and subtypes distinguished^{5,102–104}, as detailed in Figure 3A & 3B. This approach provides an inherent means towards signature-based diagnostics via data processing methods to reduce complexity in data representation. LDA models, although simple in nature, have demonstrated their utility as a diagnostic tool, and have shown promise in the development of signature-based diagnostics.

t-SNE or t-distributed stochastic neighbor embedding is a non-linear dimensionality reduction technique allowing for visualization of otherwise highly dimensional datasets¹⁰⁵. Two important considerations within t-SNE is neighbors and perplexity. Neighbors refers to points closest to or *neighboring* each point within the high dimensional dataset. For each point, t-SNE models the probability distribution of neighbors and aims to find a mapping to a lower dimensional space whilst minimizing differences between the two distributions. Perplexity can be considered a measure of neighbors which are considered when mapping the high dimensional distribution onto the lower dimensional distribution. In terms of each point within the highly dimensional distribution, controlling perplexity allows for concern of either the local or global environment around that point in turn affecting clustering efficiency and accuracy. In biological research, t-SNE has been a powerful tool for visualization of highly-dimensional biological datasets. Recently, it has been explored within the diagnostic field due to the ability to cluster exosome groups by population, elucidating tumor-derived subtypes⁵¹, results of which are shown in Figure 4A.

2.1.2 Neural and deep networks—Neural networks, including deep networks, are another intriguing class of machine learning algorithms which have gained considerable attention within healthcare. These models consist of multiple layers, each containing nodes

that are interlinked to neighboring layers, simulating neurons within the human brain¹⁰⁶. Connections between nodes of each layer have associated weights, which can alter the ‘signal’ from the previous node. The node itself applies an activation function to the weighted sum of all incoming nodes and corresponding connections, allowing for it to propagate to the next layer or stop based on a threshold of the weighted sum¹⁰⁶. Using this approach and training data, the model can ‘learn’ from the input data by tuning the weights to best classify the data. Such algorithms perform exceptionally well with image classification tasks, and have been readily integrated with image collection frameworks^{107,108}, shown in Figure 3C. With continued advancements to both device design and detection, these models may soon be applied to diagnostic platforms for big data processing. The challenge to deep network implementation is the overhead of preliminary data needed and the computational expense to train the models. Moreover, deep networks are considered ‘black-box’ approaches, meaning that the underlying features or facets the model attends to are typically not known. For molecular diagnostics, the ability to identify the underlying features could elucidate signatures beneficial for diagnostic applications.

2.2 Emerging techniques in machine learning:

Image acquisition is a common data collection step for numerous diagnostic platforms. Within the clinic, imaging can be completed with various tools, used directly for analysis, and inform diagnoses¹⁰⁹. In molecular-based diagnostics however, image acquisition typically occurs following isolation and detection of target biomarkers and is used for further processing or quantification. The image processing toolkit has been applied to clinical diagnostics and allows users to automate workflows for highly reproducible and ‘open-box’ algorithms. In other words, each step can be visualized and quantified. Such steps are highly replicable, and workflows can be created without the need for model training. The limitation to such frameworks is the need to understand the underlying data and its relationships, which can be highly complex with these large amounts of data. Image processing methods integrated with machine learning methods has shown great promise within diagnostics and molecular profiling. Though still early in development, continued efforts will likely allow for fully integrated systems for clinical diagnostics. In recent literature many microfluidic technologies have started to integrate advanced data analytics towards diagnostic development.

3. Microfluidic Technologies Integrated with Advanced Data Analytics

Integrated microfluidic technologies have greatly increased the amount of data available from on-chip platforms, necessitating advanced data analytics. The following sections cover microfluidic technologies coupled with data analysis techniques for molecular diagnostics, organized by target biomarker.

3.1 Exosomes and Microvesicles

3.1.2 Extravesicular markers—Extracellular vesicles such as exosomes and microvesicles have been largely explored as target biomarkers due to their genetic cargo and targetable surface markers. An early study in this field from the Lee and Weissleder groups utilized micro- nuclear magnetic resonance (micro-NMR) for the analysis of circulating

microvesicles¹¹⁰. This pioneering study combined immunomagnetic isolation of vesicles with micro-NMR, which allowed for microvesicle surface protein profiling directly from patient samples. In this work, they observed a discriminatory signature via relative surface protein expression (EGFR, EGFRvIII, PDPN, and IDH1 R123H), which allowed for differentiation between glioblastoma derived and host cell derived microvesicles. This four-protein signature was applied for the diagnosis of glioblastoma in clinical patient samples with an accuracy of >90%. This combinatorial approach surpassed that of any single marker alone (<76%). More recently, an integrated magneto-electrochemical sensor (iMEX) was developed by the Lee and Weissleder groups to simplify device fabrication⁹¹, highlighted in Figure 5B. Following immunomagnetic capture, this sensor, which contained a parallel electrode array of 8-electrodes, allowed for multiplex profiling of exosomal surface proteins. In a clinical setting, profiling of exosome surface proteins (EpCAM and CD24) was completed directly from the plasma of ovarian cancer patients within 1 hour using a 10 μ L sample. When compared to standard profiling methods (i.e. ELISA), this device showed high correlation ($R^2 = 0.931$) and reduced limit of detection by approximately three orders of magnitude (3×10^4 exosomes, iMEX vs. 3×10^7 , ELISA)⁹¹. Though scalability of device wasn't addressed in either study, promising primary data was presented highlighting the potential benefit of a signature-based diagnostic approach.

Immunoaffinity and immunomagnetic methods are often utilized to target overexpressed surface proteins associated with cancer-specific populations of microvesicles and exosomes. A recent study from Liu et al. utilized thermophoretic aggregation of fluorescent aptamer tagged exosomes to enable exosome isolation and profiling, the workflow and results of which is shown in Figure 5C. Seven aptamers targeting various exosomal surface proteins were introduced, and following thermophoretic aggregation, relative fluorescence intensity was used as a proxy for relative expression of each marker. Diagnostic signatures for multiple cancer subtypes (lymphoma, breast, liver, lung, ovarian, and prostate) were developed based on this approach and allowed for diagnosis of each cancer subtype⁵. A two-stage LDA classifier was coupled with this platform to develop a multi-class cancer classification model. The first classifier utilized an unweighted sum of the levels of the seven markers as the input. This unweighted sum, known as the SUM signature, was used to distinguish between cancer and healthy patients by overall expression levels. The second classifier used the individuals predicted by the first classifier to be cancer patients and aimed distinguished between the seven cancer subtypes using inputs of gender and the seven surface markers. In the second stage 15 separate LDA algorithms were created to match the number of potential cancer pairs. To achieve a single classification from these separate algorithms a max-wins voting strategy was implemented in which each classifier votes on the class and the class with the max number of votes is considered the classification. When validated within a clinical cohort, an accuracy of 99% for cancer/healthy and 68% for various cancer subtypes was shown⁵. Equally important, with this platform device fabrication is easily reproducible and scalable and required external equipment can likely be miniaturized, beneficial for point-of-care and clinical applications alike.

Plasmonics is another widely employed approach for profiling of extracellular markers. Plasmonic detection relies on the unique optical properties that occur at the interface of a metal and dielectric¹¹¹. This technique often involves the development of plasmonic

metasurfaces or the incorporation of novel plasmonic particles as capture targets. For example, Im et al. developed a nano-plasmonic sensor consisting of a periodic nanohole array, called the nPLEX assay, for profiling of exosome surface proteins^{112,113}, shown in Figure 5A. When exosomes bind near the nanoholes via immunoaffinity interactions, a characteristic change in the resonance peak in the transmission spectra can be observed and measured. This resonance peak shift can be correlated to the concentration of bound extracellular vesicles as a function of expressed surface proteins. The group applied this platform to the diagnosis of pancreatic ductal adenocarcinoma (PDAC) by targeting five surface proteins, namely EFGR, EpCAM, MUC1, WNT2, and GPC1. In this study an unweighted sum of the markers levels was used as a diagnostic metric to classify between healthy and PDAC patients. When combined and evaluated within a clinical cohort, this biomarker panel reported a sensitivity of 86%, a specificity of 81%, and an accuracy of 84%¹¹³. By utilizing this multi-biomarker signature, this platform surpassed the diagnostic performance of any single marker within this signature alone¹¹³. Although throughput and parallelization can be improved via nanohole arrays, sample purification prior to analysis is required and device fabrication is complex. Nonetheless, plasmonic detection platforms are extremely sensitive, making them promising diagnostic platforms if fabrication challenges can be overcome and allow for scalability.

Groups have begun to couple image processing techniques with plasmonic detection platforms for on-chip analysis by taking advantage of plasmonic light scattering or changes in absorption. For example, Liang et al. developed a nanoplasmon-scattering (nPES) assay, where plasmon coupling between gold nanorods and spheres was used as a means for detection⁸⁹. When both the nanorod and nanosphere were within <200 nm of one another (approximately the size of an exosome), their plasmons coupled, which resulted in a merging of the characteristic resonance peak and increased scattering intensity. Due to this increased scattering intensity, darkfield images could be taken to quantify captured exosomes. Using this methodology, a metric called the area ratio was developed to denote the number of tagged exosomes per unit area⁸⁹. In a clinical setting, area ratios were compared for normal, pancreatitis and pancreatic cancer patients. Distinct differences were observed between the three populations, demonstrating the feasibility of this simple yet powerful diagnostic tool⁸⁹. Using this methodology, a diagnostic signature could easily be developed by selecting multiple surface markers for analysis. Moreover, this technology enabled sensitive plasmonic detection overcoming the challenge of fabrication through integrating engineered plasmonic nanoparticles on a readily reproducible chip. Though machine learning for classification wasn't performed within this study, the heatmap-based output provided allows for its integration as completed within previous studies.

Image processing techniques have been employed in other systems to allow for on-chip analysis. The use of image processing has the potential to simplify and automate diagnostic workflows. In a small cohort of such studies, single exosomes were captured using a microfluidic device and detected using image processing techniques^{51,103,114}, shown in Figure 4A & 4B. This device used repeated stain and imaging cycles that targeted various exosomal surface markers to profile the distribution of surface proteins. By overlaying images, a map of the different surface markers and their relative expression was developed. Then, using t-SNE analysis for dimensionality reduction this highly dimensional data set of

the various profiled cell lines was mapped onto a 2D space whilst preserving clustering, to allow for visualization. This approach has recently been utilized by Chen et al. in an image processing and machine learning integrated framework¹⁰³. Their system utilizes a technique called DNA points accumulation for imaging in nanoscale topography (DNA-PAINT) to profile surface proteins at the single exosome level. Following imaging, relative expression of various exosomal proteins was determined and the resulting levels used as inputs for a machine learning classifier. The classifier consisted of a two-stage LDA framework, the first of which stratified healthy and cancer patients and the second of which classified between pancreatic and breast cancer. This approach allowed for near full integration of detection with downstream data analysis, and the resultant model demonstrated 100% classification accuracy between pancreatic and breast cancer patients¹⁰³. Due to the potential for high levels of diagnostic accuracy and plethora of surface markers for detection as exemplified within these previous studies, extravesicular markers have been widely explored within diagnostic platforms.

3.1.2 Intravesicular markers—Intravesicular markers of extracellular vesicles are a broad class of markers including proteins, and nucleic acids. Within extracellular vesicles, work has primarily focused on profiling of miRNAs and mRNA, though intravesicular proteins have been a recent addition to this biomarker suite¹¹⁵. An early study in this field utilized a 2-stage microfluidic capture system for exosome profiling¹⁴. The first stage captured exosomes using immunomagnetic affinity-based methods. Following on-chip lysis, a second immunomagnetic stage performed *in situ* protein capture and analysis. With this device, profiling of both surface markers and intravesicular markers was successfully shown, elucidating differences between biomarkers derived from cancer cells and those from healthy ones¹⁴.

A number of papers from Ko et al. have coupled micro-scale immunomagnetic isolation and lysis of tumor-derived exosomes with machine learning classifiers to develop biomarker signatures for disease. In these devices, overexpressed surface proteins unique to the disease of interest are tagged, and exosomes are isolated upon passing through a micro-/nano- pore filter where they are immobilized and lysed. Regarding device fabrication, these studies employed laser micromachining to create laminate sheet microfluidic devices rather than using traditional soft lithography allowing for scalability. Following extraction of mRNA¹⁰² or miRNA¹¹⁶ and quantitative polymerase chain reaction (qPCR), LDA was implemented to generate a predictive model based on a signature of exosomal mRNA or miRNA, as shown in Figure 3A. In these studies, LDA was utilized to find a panel of biomarkers which optimally classified various known disease states. With this information biomarker panels could be developed and investigated for their diagnostic potential. In the case of mRNA, the resulting panel achieved full differentiation between mice with malignant tumors, mice with premalignant lesions, and healthy mice¹⁰². The same platform was used for miRNA analysis, where a signature of eleven exosomal miRNA was able to distinguish between healthy mice, those with pancreatic ductal adenocarcinoma (PDAC), and those with precancerous lesions with 88% classification accuracy in a three-class model¹¹⁶.

In addition to the development of cancer diagnostic platforms, this group also explored the feasibility of diagnosing traumatic brain injury using this signature-based diagnostic

approach. In this case the classifier developed, which followed suit to that previously discussed, had an accuracy of 99% in classifying injured mice versus healthy mice. Results from these studies established miRNA signatures distinguishing various states of the injury including intensity, elapsed time, and presence of prior injury^{9,32}. Work by this group successfully demonstrated the feasibility of coupling micro-scale platforms to machine learning models to accurately classify specific disease states from biological data sets, representing one of the first studies in this field. Shao et al. created a microchip platform for the analysis of exosomal mRNA within tumor-derived exosomes in glioblastoma patients¹⁰. This platform, named the iMER (immune-magnetic exosomal RNA), integrated immunomagnetic isolation of target glioblastoma-derived exosomes with on-chip RNA isolation and on-chip qPCR for quantification. This tool was also used to explore efficacy of clinical interventions by measuring exosomal mRNA levels following treatment. The expression of several mRNA sequences were found to correlate with treatment¹⁰.

Though immunoaffinity and immunomagnetic approaches allow for highly specific separation and multiplexing via parallelization, the use of antibody-antigen interactions and magnetic particles as intermediaries can complicate detection and analysis. As a result, groups have investigated means to elute samples following on-chip immunoaffinity capture. The biomarker detection, quantification, and analysis can then be done after this sample elution step. Reátegui et al. integrated a thermally responsive substrate functionalized with multiple antibodies (EGFR, EFGRvIII, podoplanin, PDGFR) to target tumor derived exosomes from glioblastoma patients¹¹⁷. Once captured, by increasing the temperature to physiological temperature (37°C) the thermally responsive substrate dissolved and released captured exosomes and allowing for off-chip analysis. This platform further integrated a micro-scale geometric herringbone structure to promote turbulence in an otherwise laminar-driven system, resulting in increased capture efficiency¹¹⁷. The results demonstrated a promising means towards rapid on-chip surface protein profiling as well as the option for downstream quantitative mRNA analysis. For this work, the gene signature for multiple glioblastoma subtypes was generated using online databases. After the exosomes were isolated and studied downstream, unsupervised clustering analysis was used to analyze gene signatures to confirm capture of tumor-derived subsets and provide a potential means towards diagnosis. In this study, the geometry of the device was an assisting modality to increase capture efficiency. In some cases, geometry alone can provide a means towards capture of markers via ‘trapping’¹¹⁸ of target biomolecules. Though trapping has been largely explored for cell capture, has found limited use for extracellular vesicles due to their small size, which poses a challenge on the scale of microchannels. Nonetheless, integrated technologies have allowed for isolation and detection of intravesicular markers towards full-fledged diagnostic development.

3.2 Cells

3.2.1 Extracellular markers—Similar to extracellular vesicles, cells contain both extracellular surface proteins as well as intracellular molecular cargo such as mRNA, DNA, and proteins available for analysis. Circulating tumor cells are a subset of cells that shed from the primary cancerous tumors into the bloodstream where they can be captured and analyzed to inform diagnoses^{119,120}. An early study in this field utilized micro-NMR to

analyze cells from fine needle aspirates of patients with suspected intraabdominal tumors³. A single fine-needle aspirate was needed to provide sufficient cells, which were immunomagnetically tagged. Next, multiple surface proteins were profiled on-chip using micro-NMR. This profiling was performed on 50 patients to develop a four protein biomarker signature (EGFR, HER2, EpCAM, and MUC-1), which demonstrated 96% accuracy in cancer diagnosis³. To confirm the clinical viability of this protein signature, an independent test cohort was run, which correctly diagnosed an additional 20 patients with 100% accuracy. This early hallmark study represents the benefits that a signature-based diagnostic approach can provide. That said, this early study lacked pervasive integration of data analytics and failed to address means to scale production of the microchip, both of which would enable clinical implementation.

The ability to easily visualize and image cells following capture provides the potential for rapid image-based data analysis techniques including deep learning. A recent example of the use of deep learning for breast cancer profiling in point-of-care settings was demonstrated by Min et al. This group developed an integrated platform called AIDA (artificial intelligence diffraction analysis), which utilizes a microfluidic device to capture and stain immunolabeled cancer cells¹⁰⁸. After immunostaining, an index matched solution was applied to allow for imaging of diffraction patterns of both stained and unstained cells. Due to the large number of cells available in a single acquisition (~2,000), two deep learning algorithms were integrated to 1) detect cells, and 2) identify stained cancer cells and quantify expression levels of targeted proteins via color differences. The first deep learning algorithm achieved an accuracy of 99% in cell detection. The second algorithm achieved 90.2% accuracy in color classification, enabling differentiation between four breast cancer subtypes. Using these algorithms, analysis of clinical fine needle aspirate samples for two breast cancer patients confirmed the viability of a previously reported four-protein signature of EpCAM, EGFR, HER-2, and MUC-1 for epithelial cancer¹⁰⁸. Beneficial for clinical integration, during the development of this device emphasis was placed on low-cost materials and scalable fabrication for implementation in low-resource settings.

A second and similar platform by Im et al. called CEM (contrast enhanced microholography) profiled B cells via a capture and staining methodology for multiple surface markers¹²¹. In this platform, a microfluidic device was employed to capture B cells from fine needle aspirates of mass lesions suspected of lymphoma. Once captured, immunostaining of various surface proteins (CD19/20, Ki67, κ and λ light chains), allowed for quantification of expression following analysis. In this study, a deep learning algorithm was integrated to identify stained B cells following capture. This model utilized a convolutional neural network which was trained on >5,000 holographic cellular lymphoma images to detect B cells directly from holograms. Following analysis, the group was able to calculate the B cell size and quantify expression by degree of staining, ultimately to inform a diagnosis of lymphoma. In terms of classification accuracy for the clinical cohort tested, the CEM device had an accuracy of 95%, a sensitivity of 91% and a specificity of 100% for lymphoma diagnosis¹²¹, the workflow of which is highlighted in Figure 3D. As in the case of extravascular markers, extracellular markers are equally common targets due to the ease of access, targetability, and promising diagnostic accuracy.

3.2.2 Intracellular markers—Compared to extracellular vesicles, cells allow for intracellular protein and nucleic acid profiling due to the abundance and accessibility of intracellular molecular cargo. A study from Sundah et al. developed a barcoding scheme (called STAMP, sequence-topology assembly for multiplexed profiling) for highly multiplexed profiling of subcellular protein expression and distribution¹²², highlighted in Figure 6B. This technology hinges on tetrahedral DNA nanostructures to target unique intracellular proteins. Due to the unique DNA sequences within each tetrahedral and the number of tetrahedral conformations possible, there is a theoretical maximum of $>10^9$ unique barcodes, each of which can encode for a distinct protein¹²². Using a microfluidic device, a serpentine channel was used to enhance cell interactions with the DNA nanostructures after which a porous membrane helped to enrich the targeted cells. Following on-chip qPCR analysis and normalization, relative expression of cellular components was elucidated via fluorescence tags¹²². Using this approach, the group was able to develop molecular signatures indicative of cancer and demonstrated potential molecular markers of disease aggressiveness. When applied within a clinical cohort of cells from breast fine needle aspirate (FNA) biopsies, STAMP profiling achieved an accuracy of 94.29% when compared to clinical pathology results using a linear regression model for classification¹²². This model was built using only on-chip STAMP measurements as predictor variables and the categorical FNA status as the outcome variable¹²². Due to the levels of multiplexing possible with this approach, highly specific diagnostic signatures could be developed towards diagnostic uses. In contrast to profiling of extracellular markers, intracellular markers provide the opportunity to take advantage of traditional downstream -omic profiling methods. This allows for large amounts of data to be incorporated into the diagnostic workflow, benefiting clinical decision making.

Geometric design optimization has been used to separate cells by their physical properties and size. Sarioglu et al. created a microfluidic chip named the Cluster-Chip for the capture of circulating tumor cell clusters, associated with tumor metastasis¹¹⁸, shown in Figure 6A. This device utilized a repeating triangular trap pattern, which took advantage of the strength of cell-cell junctions for capture of these important clusters. Moreover, by reversing the flow within the device and increasing flow rates, captured clusters could be eluted for off-chip downstream immunocytochemical and molecular analysis¹¹⁸. Following elution from the chip, expression levels of transcripts from various cells including CTCs, macrophages, T cells, B cells, natural killer cells, and others, elucidated key expression profiles of metastatic breast cancer. This step of developing maps of over/under-expressed markers is a key step towards generating biomarker signatures for diagnostic applications. Though this paper didn't implement any advanced data analysis tools to aid in the diagnostic process, the nondestructive nature of CTC capture and the ability for either on- or off-chip analysis provides an ideal platform for integration. Such geometry-driven platforms typically require careful upfront design and fabrication yet can simplify working operation and greatly improve device scalability because they do not require external chemical or biological reagents.

In contrast to extracellular vesicles which in prior work have mostly been separated by host-cell type, there exist numerous specialized cell types such as immune cells which can be

profiled. Recent work from Reyes et al. has elucidated a means to separate peripheral blood mononuclear cells (PBMCs) and profile intracellular contents of multiple immune cell subsets (T cells, CD4⁺ T cells, CD8⁺ T cells, B cells, and monocytes) towards identifying a molecular disease signature for systemic lupus erythematosus (SLE)⁶⁴. This microfluidic platform incorporated both negative and positive cell isolation via immunomagnetic approaches prior to cell lysis, RNA capture, cDNA synthesis and finally amplification on-chip for RNA-seq library construction. Using this integrated device, full RNA sequencing was performed on isolated immune cell lysates within a single microfluidic chip for four markers (CD4, CD8, CD19 and CD14). Within a clinical cohort, profiling of CD4⁺ T cells, CD8⁺ T cells, B cells and CD14⁺ monocytes from five SLE patients was undertaken to highlight disease signatures of SLE⁶⁴. The signature utilized was a 37-gene expression signature for SLE previously reported, which this study used to create a signature score by summing the transcript-per-million counts for each gene⁶⁴, akin to the SUM signature discussed previously. Using this signature score within the clinical cohort showed promising differentiation between SLE patient groups and controls for B cells highlighting its potential as a SLE diagnostic panel. Intracellular molecules such as those discussed within this section represent a promising class of biomarkers within diagnostic systems, due to the abundance and heterogeneity across cell types.

3.3 Emerging Trends:

3.3.1 Device development: Previous sections have covered the most commonly profiled biomolecules using micro-scale capture and detection platforms. There are also a number of platforms in development that could be relevant to this emerging field of coupled multiplexed biomarker detection and advanced data processing methods. The Zhang research group has developed a host of technologies utilizing plasmonic and immunomagnetic approaches allowing for on-chip multiplexed circulating tumor cell detection, highlights of which are shown in Figure 7A. Through integration of external magnets and labeling circulating tumor cells using magnetic beads, separation directly from patient whole blood samples has been achieved, reducing need for sample purification. These magnetic beads can be multifunctional, and have even been used to induce 'squashing' of circulating tumor cells with the external magnets¹²³. This magnetic squashing approach coupled with a plasmonic gold chip (pGold) for detection allows for a single chip workflow via plasmonic-enhanced near-infrared (NIR) fluorescence¹²⁴. Such an approach has allowed for highly sensitivity on-chip detection of CTCs with an overall capture efficiency of ~84% and limit of detection of ~1 cell/mL whole blood¹²³.

More sensitive and highly controlled plasmonic materials, including nanorods and metasurfaces have also been a focus. A study from Tadmety et al. designed and developed peptide nucleic acid probes on plasmonic nanorods to detect point mutations in circulating tumor DNA⁴⁰. This study undertook optimization of both peptide nucleic acid probes and DNA hybridization simulations to increase specificity for point mutation detection⁴⁰. This approach could potentially further benefit from arraying of these particles, to improve sensitivity through plasmonic coupling and provide a means towards multiplexed detection. Towards this, Tadmety et al. developed a means to assemble nanorods via readily fabricated

nanowrinkled thin films⁴⁰. This study resents a unique approach which forgoes the need for complex fabrication procedures commonly associated with plasmonic detection.

Low-cost and highly scalable device development has also been an area of growing interest, targeting low-resource areas and further enabling clinical integration. Of these devices, both materials and fabrication mechanisms are often altered to avoid the need for common yet complex and resource intensive methods often implemented. One notable example includes a paper-based QR-code barcoding assay¹²⁵. This device is produced through single-step wax printing and is accompanied by a QR-barcode which allows for a simple analysis workflow¹²⁵, shown in Figure 7. Using this assay, analyte quantification can be ascertained based on processing from a cell phone image. The simplistic fabrication and operating nature of this device means that there is large potential to multiplex with distinct QR-codes for analysis. Moreover, due to the paper-based wax printing used in its fabrication, fabrication can be readily scaled for clinical use.

3.3.1 Integrated data analytics: On-chip molecular barcoding has been an emerging trend in the field, which takes advantage of innovative device design to probe biomolecules at their natural resonance frequencies. This group of promising barcode sensors, in which the Altug group has pioneered the development, is of nanophotonic metasurfaces^{90,126,127}. These platforms rely on highly precise fabrication procedures and modalities (e.g. E-beam lithography) and allow for unprecedented levels of sensitivity. An early study in this field explored the development of a pixelated dielectric metasurface with a repeating unit structure whose resonance frequency was tuned over the mid-IR range¹²⁶, shown in Figure 7B. Upon binding of analytes, a barcode-like absorption fingerprint map was created for chemical identification due to the intrinsic IR molecular vibration of the target analyte¹²⁶. Using this approach, molecular barcoding could be achieved on-chip by the photonic metasurface. A ‘fingerprint’ was created by taking advantage of the unique vibrational modes of chemical moieties within molecules. In this work, specific molecular barcodes were developed for protein A/G, a polymer mixture of PMMA and PE, and glyphosate pesticide¹²⁶.

A more recent study from the Altug group utilized an all-dielectric metasurface coupled to hyperspectral imaging and advanced image processing techniques⁹⁰, highlighted in Figure 7C. Using the methodology, the group was able to spatially resolve spectra from millions of image pixels derived from a single image, without the need for a spectrometer. The group also developed a multi-resonance sensor, which could be used for highly parallelized, fingerprint-like analysis of target biomarkers⁹⁰. In a preliminary study, detection of IgG was demonstrated at an unprecedented sensitivity of < 3 molecules/ μm^2 . Though this area of work is still relatively new, such platforms will likely provide the foundation for the next-generation of integrated, micro-scale molecular profiling and data analysis systems given alternative fabrication mechanisms allowing for scalability can be developed.

3.3.3 Advanced biomarker analysis: With advances in device development, there has been a continual progression in biomarker analyses available. A recent emerging trend in the field has been to take a step beyond the biomarker profiling discussed in depth in this review and to instead undertake functional analysis of both cells and exosomes alike^{128–130}. Such

analyses allow for probing underlying disease mechanisms to elucidate potential molecular pathways to aid in diagnosis or treatment. Though nascent in its progression, intriguing results have been presented in literature. One notable example was published by Zhang et al., which utilized nanopatterned microchips to perform both molecular and functional analysis of tumor-derived exosomes¹²⁸. The capture layer of the chip was fabricated entirely by a largely scalable inkjet printing technique developed by the group which allows for high-resolution colloidal printing. This substrate provided robust immunocapture and was coupled with on-chip enzyme linked immunosorbent (ELISA) and proteolytic activity assays for functional analyses. With this device, the group investigated feasibility a novel potential clinical biomarker, matrix metalloproteinase-14 (MMP14) and demonstrated its application for monitoring tumor growth and metastasis in mice. For clinical validation, they coupled on-chip analysis with a LDA model for breast cancer detection and staging with an accuracy of 92.9%¹²⁸.

3.4 Industry Technologies:

Within the context of industrial and clinical integration, there are a small subsection of commercial technologies which have the potential to allow for signature-based diagnostics. This class of technologies fundamentally allows for highly multiplexed biomarker analysis, typically from purified patient blood samples¹³¹. For biomarker isolation within the clinic, centrifugation of collected patient blood samples is often the preliminary step prior to technology integration and analysis, in order to separate blood into its constituents (plasma, buffy coat, and red blood cells). From here various additional isolation techniques for specific biomolecules of interest are completed, specific for the analysis being completed and system requirements. As in the case of laboratory-based devices, many of the current devices allow only for big data generation and visualization. After this stage, integrated data analysis techniques can be separately introduced to parse through these complex datasets. Fully integrated technologies with data generation and machine learning or image processing integration are being proposed in a number of start-up companies, yet adoption is still unknown.

Of the available clinical technologies, NanoString's nCounter system and Biocartis's Idylla system are two notable examples. Both examples represent a sample in, data out set of technologies integrating biomarker isolation, detection and analysis. NanoString's nCounter system is an automated molecular barcoding technology, allowing for analysis of biomarkers (miRNA, protein, mRNA expression) from samples such as cell and tissue lysates, biomolecules such as exosomes, and formalin-fixed paraffin-embedded tissue (FFPE). In this system, due to the unique barcode technology and accompanying image processing software which enumerates these barcodes, highly multiplexed measurement is feasible^{132,133}. This parallelization in capture and detection allows for expression profiles to be generated, aiding the development in signature-based biomarker panels beneficial for diagnostic implementation. Though many of their assays are made for research use only, there exists an FDA approved assay known as the Prosigna Assay, which uses a gene signature to identify the patient's risk breast cancer recurrence^{134,135}.

Biocartis's Idylla system is a highly automated molecular technology specifically developed for use within oncology. This system allows for automation of both sample preparation as well as the data analysis and reporting, with the analysis workflow taking only 90–150 minutes¹³⁶. Analysis can be achieved from patient plasma or FFPE tissue slices alike, and samples are input into a cartridge which facilitates sample preparation and analysis. Biocartis has developed numerous multiplexed assays which access and qualitatively detect key mutations within exons and codons of various genes and oncogenes such as EGFR and KRAS¹³⁷. Unlike NanoString which doesn't require amplification of target sequences and rather utilizes the barcoding for enumeration, the Idylla system automates real-time PCR for to allow for detection of these key mutations. These technologies show how integrated biomarker quantification and analysis integration will continue to play a foundational role in research and clinical settings.

4. Conclusions:

Signature-based diagnostics have been a topic of increasing interest due to promising results in terms of diagnostic accuracy and big data integration to inform diagnoses. These technologies employ microfluidic biomarker screening technologies as a fundamental building blocks due to their low sample volume requirements, high-throughput, and ease of parallelization. Additionally, on-chip analysis can be achieved via molecular profiling enabled by multiplex detection. As a result of the increasing parallelization of these systems, there is an increase in data output. This premium in data has catapulted the use of advanced data techniques such as machine learning and image processing to aid in the diagnostic process. Given the emergence of these new integrated hardware and software systems, there exists a continual need to keep clinical utility at the forefront of the device and analysis development framework.

Although this field is relatively new, promising results have already been demonstrated in the literature. There now exist a number of multiplexed device designs and concepts, tailored towards an ever-increasing library of biomarkers. This review highlighted many of the recent, promising devices that integrate advanced data processing techniques. Though significant progress has been made, there exists numerous challenges that remain on both the device integration and data analysis sides of such technologies. Namely, improvements are needed in device design to allow for increased parallelization and flexibility in terms of biomarkers profiled and analysis completed. At present, many devices focus on capturing multiple subpopulations of the same molecule. An interesting, and previously unreported alternative, would be the design, capture and profiling of multiple molecules (e.g. CTCs and exosomes) to develop a “multi-biomarker” signature. Integration of multiple molecules could potentially enhance diagnostic potential whilst increasing data available for diagnoses and feasibility for advanced data analysis integration.

Concerning data analysis methodologies, though now more widespread and accessible than ever, there needs to be a facile means to integrate and explore the potential benefits of this integration. Molecular diagnostics provide large amount of data available to harness for diagnostic purposes. Continual advances in existing device designs and the development of novel new systems allow for highly parallelized analysis with ever-decreasing sample-to-

answer times. With these advances, there is an increasing need to incorporate advanced data analysis techniques to synthesize these complex biological datasets to gather meaningful and actionable data beneficial for diagnoses. To increase feasibility and address benefits and limitations of various systems, standardization and guides as to the data fit for each data analysis technique and the means for integration would be beneficial. Likewise, the ability to corroborate platforms able to gather this plethora of data and experts in data analysis is an increasing necessity.

In addition to advances in device design and data analysis methodologies by themselves, seamless merging between micro-scale devices and data processing techniques is needed. Moving forward, the development of an integrated software platform for universal use could enable uniform downstream data processing workflow for numerous diagnostic platforms. Nonetheless, promising results have already been reported using signature-based diagnostic platforms. With continual advances, these integrated platforms have the potential to improve disease diagnostics, make contributions to personalized medicine, and improve patient outcomes.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

5. References:

1. Ngo A, Gandhi P, Miller WG. Frequency that Laboratory Tests Influence Medical Decisions. *J Appl Lab Med An AACC Publ.* 2017;1(4):410–414. doi:10.1373/jalm.2016.021634
2. Borrebaeck CAK. Precision diagnostics: moving towards protein biomarker signatures of clinical utility in cancer. *Nat Rev Cancer.* 2017;17(3):199–204. doi:10.1038/nrc.2016.153 [PubMed: 28154374]
3. Haun JB, Castro CM, Wang R, et al. Micro-NMR for rapid molecular analysis of human tumor samples. *Sci Transl Med.* 2011;3(71):71ra16. doi:10.1126/scitranslmed.3002048
4. Adela R, Reddy PNC, Ghosh TS, et al. Serum protein signature of coronary artery disease in type 2 diabetes mellitus. *J Transl Med.* 2019;17(1):17. doi:10.1186/s12967-018-1755-5 [PubMed: 30674322]
5. Liu C, Zhao J, Tian F, et al. Low-cost thermophoretic profiling of extracellular-vesicle surface proteins for the early detection and classification of cancers. *Nat Biomed Eng.* 2019;3(3):183–193. doi:10.1038/s41551-018-0343-6 [PubMed: 30948809]
6. Zhang H, Zhu M, Shan X, et al. A panel of seven-miRNA signature in plasma as potential biomarker for colorectal cancer diagnosis. *Gene.* 2019;687:246–254. doi:10.1016/j.gene.2018.11.055 [PubMed: 30458288]
7. Chen M, Xu R, Rai A, et al. Distinct shed microvesicle and exosome microRNA signatures reveal diagnostic markers for colorectal cancer. Ahmad A, ed. *PLoS One.* 2019;14(1):e0210003. doi:10.1371/journal.pone.0210003 [PubMed: 30608951]
8. Garcia-Contreras M, Shah SH, Tamayo A, et al. Plasma-derived exosome characterization reveals a distinct microRNA signature in long duration Type 1 diabetes. *Sci Rep.* 2017;7(1):5998. doi:10.1038/s41598-017-05787-y [PubMed: 28729721]
9. Ko J, Hemphill MA, Yang Z, et al. Diagnosis of traumatic brain injury using miRNA signatures in nanomagnetically isolated brain-derived extracellular vesicles. *Lab Chip.* 2018. doi:10.1039/C8LC00672E
10. Shao H, Chung J, Lee K, et al. Chip-based analysis of exosomal mRNA mediating drug resistance in glioblastoma. *Nat Commun.* 2015;6(5):2–10. doi:10.1038/ncomms7999

11. Nguyen AH, Sim SJ. Nanoplasmonic biosensor: Detection and amplification of dual bio-signatures of circulating tumor DNA. *Biosens Bioelectron.* 2015;67:443–449. doi:10.1016/J.BIOS.2014.09.003 [PubMed: 25220802]
12. Hisey CL, Dorayappan KDP, Cohn DE, Selvendiran K, Hansford DJ. Microfluidic affinity separation chip for selective capture and release of label-free ovarian cancer exosomes. *Lab Chip.* 2018;3144–3153. doi:10.1039/C8LC00834E [PubMed: 30191215]
13. San Lucas FA, Allenson K, Bernard V, et al. Minimally invasive genomic and transcriptomic profiling of visceral cancers by next-generation sequencing of circulating exosomes. *Ann Oncol.* 2016;27(4):635–641. doi:10.1093/annonc/mdv604 [PubMed: 26681674]
14. He M, Crow J, Roth M, Zeng Y, Godwin AK. Integrated immunoisolation and protein analysis of circulating exosomes using microfluidic technology. *Lab Chip.* 2014;14(19):3773–3780. doi:10.1039/c4lc00662c [PubMed: 25099143]
15. Chen C, Zhao X-L, Li Z-H, Zhu Z-G, Qian S-H, Flewitt AJ. Current and Emerging Technology for Continuous Glucose Monitoring. *Sensors (Basel).* 2017;17(1). doi:10.3390/s17010182
16. Borel F, Konstantinova P, Jansen PLM. Diagnostic and therapeutic potential of miRNA signatures in patients with hepatocellular carcinoma. *J Hepatol.* 2012;56(6):1371–1383. doi:10.1016/J.JHEP.2011.11.026 [PubMed: 22314424]
17. Wei W, Shin YS, Ma C, et al. Microchip platforms for multiplex single-cell functional proteomics with applications to immunology and cancer research. *Genome Med.* 2013;5(8):75. doi:10.1186/gm479 [PubMed: 23998271]
18. Kantamneni H, Zevon M, Donzanti MJ, et al. Surveillance nanotechnology for multi-organ cancer metastases. *Nat Biomed Eng.* 2017;1(12):993–1003. doi:10.1038/s41551-017-0167-9 [PubMed: 29531851]
19. Wan Y, Cheng G, Liu X, et al. Rapid magnetic isolation of extracellular vesicles via lipid-based nanoprobe. *Nat Biomed Eng.* 2017;1(4):1–11. doi:10.1038/s41551-017-0058
20. Panagopoulou M, Karaglanis M, Balgkouranidou I, et al. Circulating cell-free DNA in breast cancer: size profiling, levels, and methylation patterns lead to prognostic and predictive classifiers. *Oncogene.* 2019;38(18):3387–3401. doi:10.1038/s41388-018-0660-y [PubMed: 30643192]
21. Whiteside TL. The emerging role of plasma exosomes in diagnosis, prognosis and therapies of patients with cancer. *Wspolczesna Onkol.* 2017;2(1A):38–40. doi:10.5114/wo.2018.73882
22. Zhao Z, Fan X, Yang L, et al. The identification of a common different gene expression signature in patients with colorectal cancer. *Math Biosci Eng.* 2019;16(4):2942–2958. doi:10.3934/mbe.2019145 [PubMed: 31137244]
23. Wu M, Huang PH, Zhang R, et al. Circulating Tumor Cell Phenotyping via High-Throughput Acoustic Separation. *Small.* 2018;14(32):1–10. doi:10.1002/sml.201801131
24. Carmichael J, Hayashi C, Huang X, et al. Label-free characterization of exosome via surface enhanced Raman spectroscopy for the early detection of pancreatic cancer. *Nanomedicine Nanotechnology, Biol Med.* 2019;16:88–96. doi:10.1016/J.NANO.2018.11.008
25. Fehlbaum-Beurdeley P, Jarrige-Le Prado AC, Pallares D, et al. Toward an Alzheimer's disease diagnosis via high-resolution blood gene expression. *Alzheimer's Dement.* 2010;6(1):25–38. doi:10.1016/J.JALZ.2009.07.001 [PubMed: 20129318]
26. Lee S, Mankhong S, Kang J-H. Extracellular Vesicle as a Source of Alzheimer's Biomarkers: Opportunities and Challenges. *Int J Mol Sci.* 2019;20(7):1728. doi:10.3390/ijms20071728
27. Naveed M, Mubeen S, Khan A, Ibrahim S, Meer B. Plasma Biomarkers: Potent Screeners of Alzheimer's Disease. *Am J Alzheimer's Dis Other Dementias®.* 5 2019:153331751984823. doi:10.1177/1533317519848239
28. Kumar P, Dezso Z, MacKenzie C, et al. Circulating miRNA Biomarkers for Alzheimer's Disease. Hill AF, ed. *PLoS One.* 2013;8(7):e69807. doi:10.1371/journal.pone.0069807 [PubMed: 23922807]
29. Habartová L, Hrušková K, Syslová K, et al. Blood-based molecular signature of Alzheimer's disease via spectroscopy and metabolomics. *Clin Biochem.* 4 2019. doi:10.1016/J.CLINBIOCHEM.2019.04.004

30. Rahman MR, Islam T, Shahjaman M, Quinn JMW, Holsinger RMD, Moni MA. Identification of common molecular biomarker signatures in blood and brain of Alzheimer's disease. *bioRxiv*. 1 2019;482828. doi:10.1101/482828
31. Pizarro C, Esteban-Díez I, Espinosa M, Rodríguez-Royo F, González-Sáiz J-M. An NMR-based lipidomic approach to identify Parkinson's disease-stage specific lipoprotein-lipid signatures in plasma. *Analyst*. 2019;144(4):1334–1344. doi:10.1039/C8AN01778F [PubMed: 30564825]
32. Ko J, Hemphill M, Yang Z, et al. Multi-Dimensional Mapping of Brain-Derived Extracellular Vesicle MicroRNA Biomarker for Traumatic Brain Injury Diagnostics. *J Neurotrauma*. 2019;11:1–11. doi:10.1089/neu.2018.6220
33. Ren F, Gao W, Ke Z, Xu Y, Liu Y. Identification of six miRNAs serving as predictive biomarkers in coronary artery disease. *J Cell Biochem*. 2019;120(2):1932–1942. doi:10.1002/jcb.27476
34. Hao N, Zhang JXJ. Microfluidic Screening of Circulating Tumor Biomarkers toward Liquid Biopsy. *Sep Purif Rev*. 2018;47(1):19–48. doi:10.1080/15422119.2017.1320763
35. Ying L, Wang Q. Microfluidic chip-based technologies: emerging platforms for cancer diagnosis. *BMC Biotechnol*. 2013;13(1):76. doi:10.1186/1472-6750-13-76 [PubMed: 24070124]
36. Cima I, Wen Yee C, Iliescu FS, et al. Label-free isolation of circulating tumor cells in microfluidic devices: Current research and perspectives. *Biomicrofluidics*. 2013;7(1):011810. doi:10.1063/1.4780062
37. Norgeot B, Glicksberg BS, Butte AJ. A call for deep-learning healthcare. *Nat Med*. 2019;25(1):14–15supp. doi:10.1038/s41591-018-0320-3 [PubMed: 30617337]
38. Raghupathi W, Raghupathi V. Big data analytics in healthcare: promise and potential. *Heal Inf Sci Syst*. 2014;2(1):3. doi:10.1186/2047-2501-2-3
39. Wallwiener M, Riethdorf S, Hartkopf AD, et al. Serial enumeration of circulating tumor cells predicts treatment response and prognosis in metastatic breast cancer: a prospective study in 393 patients. *BMC Cancer*. 2014;14(1):512. doi:10.1186/1471-2407-14-512 [PubMed: 25015676]
40. Tadimety A, Zhang Y, Kready KM, Palinski TJ, Tsongalis GJ, Zhang JXJ. Design of peptide nucleic acid probes on plasmonic gold nanorods for detection of circulating tumor DNA point mutations. *Biosens Bioelectron*. 2019;130:236–244. doi:10.1016/j.bios.2019.01.045 [PubMed: 30769288]
41. Pan Y, Liu G, Yuan Y, Zhao J, Yang Y, Li Y. Analysis of differential gene expression profile identifies novel biomarkers for breast cancer. *Oncotarget*. 2017;8(70):114613–114625. doi:10.18632/oncotarget.23061 [PubMed: 29383106]
42. Yang Z, Wu L, Wang A, et al. DbDEMC 2.0: Updated database of differentially expressed miRNAs in human cancers. *Nucleic Acids Res*. 2017;45(D1):D812–D818. doi:10.1093/nar/gkw1079 [PubMed: 27899556]
43. Kamel HFM, Al-Amodi HSAB. Exploitation of Gene Expression and Cancer Biomarkers in Paving the Path to Era of Personalized Medicine. *Genomics Proteomics Bioinformatics*. 2017;15(4):220–235. doi:10.1016/j.gpb.2016.11.005 [PubMed: 28813639]
44. Che J, Yu V, Dhar M, et al. Classification of large circulating tumor cells isolated with ultra-high throughput microfluidic Vortex technology. *Oncotarget*. 2016;7(11):12748–12760. doi:10.18632/oncotarget.7220 [PubMed: 26863573]
45. Munz M, Baeuerle PA, Gires O. The Emerging Role of EpCAM in Cancer and Stem Cell Signaling. *Cancer Res*. 2009;69(14):5627–5629. doi:10.1158/0008-5472.CAN-09-0654 [PubMed: 19584271]
46. Basakran NS. CD44 as a potential diagnostic tumor marker. *Saudi Med J*. 2015;36(3):273–279. doi:10.15537/smj.2015.3.9622 [PubMed: 25737167]
47. Vences-Catalán F, Duault C, Kuo C-C, Rajapaksa R, Levy R, Levy S. CD81 as a tumor target. *Biochem Soc Trans*. 2017;45(2):531–535. doi:10.1042/BST20160478 [PubMed: 28408492]
48. Soung Y, Ford S, Zhang V, Chung J. Exosomes in Cancer Diagnostics. *Cancers (Basel)*. 2017;9(12):8. doi:10.3390/cancers9010008
49. Kurywchak P, Tavormina J, Kalluri R. The emerging roles of exosomes in the modulation of immune responses in cancer. *Genome Med*. 2018;10(1):1–4. doi:10.1186/s13073-018-0535-4 [PubMed: 29301565]

50. Théry C, Zitvogel L, Amigorena S. Exosomes: composition, biogenesis and function. *Nat Rev Immunol.* 2002;2(8):569–579. doi:10.1038/nri855 [PubMed: 12154376]
51. Lee K, Fraser K, Ghaddar B, et al. Multiplexed Profiling of Single Extracellular Vesicles. *ACS Nano.* 2018;12(1):494–503. doi:10.1021/acsnano.7b07060 [PubMed: 29286635]
52. Colombo M, Raposo G, Théry C. Biogenesis, Secretion, and Intercellular Interactions of Exosomes and Other Extracellular Vesicles. *Annu Rev Cell Dev Biol.* 2014;30(1):255–289. doi:10.1146/annurev-cellbio-101512-122326 [PubMed: 25288114]
53. Wang Z, Zöller M. Exosomes, metastases, and the miracle of cancer stem cell markers. *Cancer Metastasis Rev.* 4 2019:1–37. doi:10.1007/s10555-019-09793-6 [PubMed: 31297746]
54. Maia J, Caja S, Strano Moraes MC, Couto N, Costa-Silva B. Exosome-Based Cell-Cell Communication in the Tumor Microenvironment. *Front Cell Dev Biol.* 2018;6(2):1–19. doi:10.3389/fcell.2018.00018 [PubMed: 29417046]
55. Gold B, Gocke CD, Meier F, Furtado LV., Cankovic M. Do Circulating Tumor Cells, Exosomes, and Circulating Tumor Nucleic Acids Have Clinical Utility? *J Mol Diagnostics.* 2015;17(3):209–224. doi:10.1016/j.jmoldx.2015.02.001
56. Whiteside TL. The potential of tumor-derived exosomes for noninvasive cancer monitoring. *Expert Rev Mol Diagn.* 2015;15(10):1293–1310. doi:10.1586/14737159.2015.1071666 [PubMed: 26289602]
57. Zhao Z, Yang Y, Zeng Y, He M. A microfluidic ExoSearch chip for multiplexed exosome detection towards blood-based ovarian cancer diagnosis. *Lab Chip.* 2016;16(3):489–496. doi:10.1039/c5lc01117e [PubMed: 26645590]
58. Khan S, Jutzy JMS, Valenzuela MMA, et al. Plasma-Derived Exosomal Survivin, a Plausible Biomarker for Early Detection of Prostate Cancer. Li J, ed. *PLoS One.* 2012;7(10):e46737. doi:10.1371/journal.pone.0046737 [PubMed: 23091600]
59. Rupp A-K, Rupp C, Keller S, et al. Loss of EpCAM expression in breast cancer derived serum exosomes: Role of proteolytic cleavage. *Gynecol Oncol.* 2011;122(2):437–446. doi:10.1016/j.ygyno.2011.04.035 [PubMed: 21601258]
60. Schwarzenbach H, Hoon DSB, Pantel K. Cell-free nucleic acids as biomarkers in cancer patients. *Nat Rev Cancer.* 2011;11(6):426–437. doi:10.1038/nrc3066 [PubMed: 21562580]
61. Tadimety A, Closson A, Li C, Yi S, Shen T, Zhang XJ. Advances in liquid biopsy on-chip for cancer management: Technologies, biomarkers, and clinical analysis. *Crit Rev Clin Lab Sci.* 2018;55(3):1–23. doi:10.1080/10408363.2018.1425976 [PubMed: 29239240]
62. Tadimety A, Syed A, Nie Y, Long CR, Kready KM, Zhang XJ. Liquid biopsy on chip: a paradigm shift towards the understanding of cancer metastasis. *Integr Biol.* 2017;9(1):22–49. doi:10.1039/C6IB00202A
63. Ma S, Hsieh Y-P, Ma J, Lu C. Low-input and multiplexed microfluidic assay reveals epigenomic variation across cerebellum and prefrontal cortex. *Sci Adv.* 2018;4(4):eaar8187. doi:10.1126/sciadv.aar8187 [PubMed: 29675472]
64. Reyes M, Vickers D, Billman K, et al. Multiplexed enrichment and genomic profiling of peripheral blood cells reveal subset-specific immune signatures. *Sci Adv.* 2019;5(1):eaau9223. doi:10.1126/sciadv.aau9223 [PubMed: 30746468]
65. Ozawa R, Iwadata H, Toyoda H, Yamada M, Seki M. A numbering-up strategy of hydrodynamic microfluidic filters for continuous-flow high-throughput cell sorting. *Lab Chip.* 2019;19(10):1828–1837. doi:10.1039/C9LC00053D [PubMed: 30998230]
66. Ko J, Baldassano SN, Loh PL, Kording K, Litt B, Issadore D. Machine learning to detect signatures of disease in liquid biopsies-a user's guide. *Lab Chip.* 2018;18(3):395–405. doi:10.1039/c7lc00955k [PubMed: 29192299]
67. Gou T, Hu J, Zhou S, et al. A new method using machine learning for automated image analysis applied to chip-based digital assays. *Analyst.* 2019;144(10):3274–3281. doi:10.1039/C9AN00149B [PubMed: 30990486]
68. Turan B, Masuda T, Lei W, et al. A pillar-based microfluidic chip for T-cells and B-cells isolation and detection with machine learning algorithm. *ROBOMECH J.* 2018;5(1). doi:10.1186/s40648-018-0124-8

69. Burkland A, Zhang JXJ. Microfluidics-Based Organism Isolation from Whole Blood: An Emerging Tool for Bloodstream Infection Diagnosis. *Ann Biomed Eng.* 2019;47(7):1657–1674. doi:10.1007/s10439-019-02256-7 [PubMed: 30980291]
70. Contreras-Naranjo JC, Wu H-J, Ugaz VM. Microfluidics for exosome isolation and analysis: enabling liquid biopsy for personalized medicine. *Lab Chip.* 2017;17(21):3558–3577. doi:10.1039/c7lc00592j [PubMed: 28832692]
71. Cho H, Kim J, Song H, Sohn KY, Jeon M, Han K-H. Microfluidic technologies for circulating tumor cell isolation. *Analyst.* 2018;143(13):2936–2970. doi:10.1039/c7an01979c [PubMed: 29796523]
72. Sonnenberg A, Marciniak JY, Skowronski EA, et al. Dielectrophoretic isolation and detection of cancer-related circulating cell-free DNA biomarkers from blood and plasma. *Electrophoresis.* 2014;35(12–13):1828–1836. doi:10.1002/elps.201400016 [PubMed: 24723219]
73. Chan JY, Ahmad Kayani A Bin, Md Ali MA, et al. Dielectrophoresis-based microfluidic platforms for cancer diagnostics. *Biomicrofluidics.* 2018;12(1). doi:10.1063/1.5010158
74. Ding X, Peng Z, Lin S-CS, et al. Cell separation using tilted-angle standing surface acoustic waves. *Proc Natl Acad Sci U S A.* 2014;111(36):12992–12997. doi:10.1073/pnas.1413325111 [PubMed: 25157150]
75. Lu X, Martin A, Soto F, et al. Parallel Label-Free Isolation of Cancer Cells Using Arrays of Acoustic Microstreaming Traps. *Adv Mater Technol.* 2018;4(2):1800374. doi:10.1002/admt.201800374
76. Ozcelik A, Rufo J, Guo F, et al. Acoustic tweezers for the life sciences. *Nat Methods.* 2018;15(12):1021–1028. doi:10.1038/s41592-018-0222-9 [PubMed: 30478321]
77. Wu M, Chen C, Wang Z, et al. Separating extracellular vesicles and lipoproteins via acoustofluidics. *Lab Chip.* 2019;19(7):1174–1182. doi:10.1039/c8lc01134f [PubMed: 30806400]
78. Sun J, Liu C, Li M, et al. Size-based hydrodynamic rare tumor cell separation in curved microfluidic channels. *Biomicrofluidics.* 2013;7(1):1–11. doi:10.1063/1.4774311
79. Ribeiro-Samy S, Oliveira MI, Pereira-Veiga T, et al. Fast and efficient microfluidic cell filter for isolation of circulating tumor cells from unprocessed whole blood of colorectal cancer patients. *Sci Rep.* 2019;9(1):8032. doi:10.1038/s41598-019-44401-1 [PubMed: 31142796]
80. Wunsch BH, Smith JT, Gifford SM, et al. Nanoscale lateral displacement arrays for the separation of exosomes and colloids down to 20nm. *Nat Nanotechnol.* 2016;11(11):936–940. doi:10.1038/nnano.2016.134 [PubMed: 27479757]
81. Chen H A Triplet Parallelizing Spiral Microfluidic Chip for Continuous Separation of Tumor Cells. *Sci Rep.* 2018;8(1):4042. doi:10.1038/s41598-018-22348-z [PubMed: 29511230]
82. Huang LR, Cox EC, Austin RH, Sturm JC. Continuous Particle Separation Through Deterministic Lateral Displacement. *Science (80-).* 2004;304(5673):987–990. doi:10.1126/science.1094567
83. Hoshino K, Huang YY, Lane N, et al. Microchip-based immunomagnetic detection of circulating tumor cells. *Lab Chip.* 2011;11(20):3449–3457. doi:10.1039/c1lc20270g [PubMed: 21863182]
84. Huang YY, Chen P, Wu CH, et al. Screening and Molecular Analysis of Single Circulating Tumor Cells Using Micromagnet Array. *Sci Rep.* 2015;5:1–11. doi:10.1038/srep16047
85. Bhuvanendran Nair Gourikutty S, Chang C-P, Pui PD. Microfluidic immunomagnetic cell separation from whole blood. *J Chromatogr B.* 2016;1011:77–88. doi:10.1016/J.JCHROMB.2015.12.016
86. Shi W, Wang S, Maarouf A, et al. Magnetic particles assisted capture and release of rare circulating tumor cells using wavy-herringbone structured microfluidic devices. *Lab Chip.* 2017;17(19):3291–3299. doi:10.1039/c7lc00333a [PubMed: 28840927]
87. Ng E, Hoshino K, Zhang X. Microfluidic immunodetection of cancer cells via site-specific microcontact printing of antibodies on nanoporous surface. *Methods.* 2013;63(3):266–275. doi:10.1016/j.ymeth.2013.07.043 [PubMed: 24012763]
88. Wu CH, Huang YY, Chen P, et al. Versatile immunomagnetic nanocarrier platform for capturing cancer cells. *ACS Nano.* 2013;7(10):8816–8823. doi:10.1021/nn403281e [PubMed: 24016305]
89. Liang K, Liu F, Fan J, et al. Nanoplasmonic quantification of tumour-derived extracellular vesicles in plasma microsamples for diagnosis and treatment monitoring. *Nat Biomed Eng.* 2017;1(4):0021. doi:10.1038/s41551-016-0021 [PubMed: 28791195]

90. Yesilkoy F, Arvelo ER, Jahani Y, et al. Ultrasensitive hyperspectral imaging and biodetection enabled by dielectric metasurfaces. *Nat Photonics*. 2019. doi:10.1038/s41566-019-0394-6
91. Jeong S, Park J, Pathania D, Castro CM, Weissleder R, Lee H. Integrated Magneto–Electrochemical Sensor for Exosome Analysis. *ACS Nano*. 2016;10(2):1802–1809. doi:10.1021/acsnano.5b07584 [PubMed: 26808216]
92. Safaei TS, Mohamadi RM, Sargent EH, Kelley SO. In situ electrochemical ELISA for specific identification of captured cancer cells. *ACS Appl Mater Interfaces*. 2015;7(26):14165–14169. doi:10.1021/acsami.5b02404 [PubMed: 25938818]
93. Cima I, Schiess R, Wild P, et al. Cancer genetics-guided discovery of serum biomarker signatures for diagnosis and prognosis of prostate cancer. *Proc Natl Acad Sci U S A*. 2011;108(8):3342–3347. doi:10.1073/pnas.1013699108 [PubMed: 21300890]
94. Li L, Wei Y, To C, et al. Integrated Omic analysis of lung cancer reveals metabolism proteome signatures with prognostic impact. *Nat Commun*. 2014;5(1):5469. doi:10.1038/ncomms6469 [PubMed: 25429762]
95. Daemen A, Griffith OL, Heiser LM, et al. Modeling precision treatment of breast cancer. *Genome Biol*. 2013;14(10):R110. doi:10.1186/gb-2013-14-10-r110 [PubMed: 24176112]
96. Liga A, Vliegthart ADB, Oosthuyzen W, Dear JW, Kersaudy-Kerhoas M. Exosome isolation: a microfluidic road-map. *Lab Chip*. 2015;15(11):2388–2394. doi:10.1039/c5lc00240k [PubMed: 25940789]
97. Dong Y, Skelley AM, Merdek KD, et al. Microfluidics and Circulating Tumor Cells. *J Mol Diagnostics*. 2013;15(2):149–157. doi:10.1016/J.JMOLDX.2012.09.004
98. He M, Zeng Y. Microfluidic Exosome Analysis toward Liquid Biopsy for Cancer. *J Lab Autom*. 2016;21(4):599–608. doi:10.1177/2211068216651035 [PubMed: 27215792]
99. Rana A, Zhang Y, Esfandiari L. Advancements in microfluidic technologies for isolation and early detection of circulating cancer-related biomarkers. *Analyst*. 2018;143(13):2971–2991. doi:10.1039/C7AN01965C [PubMed: 29790491]
100. Dan Z, Daxiang C. Advances in isolation and detection of circulating tumor cells based on microfluidics. *Cancer Biol Med*. 2018;15(4):335. doi:10.20892/j.issn.2095-3941.2018.0256 [PubMed: 30766747]
101. Mika S, Ratsch G, Weston J, Scholkopf B, Mullers KR. Fisher discriminant analysis with kernels. In: *Neural Networks for Signal Processing IX: Proceedings of the 1999 IEEE Signal Processing Society Workshop* (Cat. No.98TH8468). IEEE;:41–48. doi:10.1109/NNSP.1999.788121
102. Ko J, Bhagwat N, Yee SS, et al. Combining Machine Learning and Nanofluidic Technology To Diagnose Pancreatic Cancer Using Exosomes. *ACS Nano*. 2017;11(11):11182–11193. doi:10.1021/acsnano.7b05503 [PubMed: 29019651]
103. Chen C, Zong S, Liu Y, et al. Profiling of Exosomal Biomarkers for Accurate Cancer Identification: Combining DNA-PAINT with Machine- Learning-Based Classification. *Small*. 9 2019;1901014. doi:10.1002/sml.201901014
104. Zhang P, Zhou X, Zeng Y. Multiplexed immunophenotyping of circulating exosomes on nano-engineered ExoProfile chip towards early diagnosis of cancer. *Chem Sci*. 2019;10(21):5495–5504. doi:10.1039/c9sc00961b [PubMed: 31293733]
105. van der Maaten L, Hinton G. Visualizing Data using t-SNE. *J Mach Learn Res*. 2008;9(11):2579–2605. <http://www.jmlr.org/papers/v9/vandermaaten08a.html>. Accessed November 13, 2019.
106. Basheer I, Hajmeer M. Artificial neural networks: fundamentals, computing, design, and application. *J Microbiol Methods*. 2000;43(1):3–31. doi:10.1016/S0167-7012(00)00201-3 [PubMed: 11084225]
107. Kim S-J, Wang C, Zhao B, et al. Deep transfer learning-based hologram classification for molecular diagnostics. *Sci Rep*. 2018;8(1):17003. doi:10.1038/s41598-018-35274-x [PubMed: 30451953]
108. Min J, Im H, Allen M, et al. Computational Optics Enables Breast Cancer Profiling in Point-of-Care Settings. *ACS Nano*. 2018;12s(9):9081–9090. doi:10.1021/acsnano.8b03029
109. Dougherty G. Image analysis in medical imaging: recent advances in selected examples. *Biomed Imaging Interv J*. 2010;6(3):e32. doi:10.2349/bij.6.3.e32 [PubMed: 21611048]

110. Shao H, Chung J, Balaj L, et al. Protein typing of circulating microvesicles allows real-time monitoring of glioblastoma therapy. *Nat Med.* 2012;18(12):1835–1840. doi:10.1038/nm.2994 [PubMed: 23142818]
111. Mayer KM, Hafner JH. Localized Surface Plasmon Resonance Sensors. *Chem Rev.* 2011;111(6):3828–3857. doi:10.1021/cr100313v [PubMed: 21648956]
112. Im H, Shao H, Park Y Il, et al. Label-free detection and molecular profiling of exosomes with a nano-plasmonic sensor. *Nat Biotechnol.* 2014;32(5):490–495. doi:10.1038/nbt.2886 [PubMed: 24752081]
113. Yang KS, Im H, Hong S, et al. Multiparametric plasma EV profiling facilitates diagnosis of pancreatic malignancy. *Sci Transl Med.* 2017;9(391):eaal3226. doi:10.1126/scitranslmed.aal3226 [PubMed: 28539469]
114. Fraser K, Jo A, Giedt J, et al. Characterization of single microvesicles in plasma from glioblastoma patients. *Neuro Oncol.* 2019;21(5):606–615. doi:10.1093/neuonc/noy187 [PubMed: 30561734]
115. Park J, Im H, Hong S, Castro CM, Weissleder R, Lee H. Analyses of Intravesicular Exosomal Proteins Using a Nano-Plasmonic System. *ACS Photonics.* 2018;5(2):487–494. doi:10.1021/acsp Photonics.7b00992 [PubMed: 29805987]
116. Ko J, Bhagwat N, Black T, et al. MiRNA profiling of magnetic nanopore-isolated extracellular vesicles for the diagnosis of pancreatic cancer. *Cancer Res.* 2018;78(13):3688–3697. doi:10.1158/0008-5472.CAN-17-3703 [PubMed: 29735554]
117. Reátegui E, Van Der Vos KE, Lai CP, et al. Engineered nanointerfaces for microfluidic isolation and molecular profiling of tumor-specific extracellular vesicles. *Nat Commun.* 2018;9(1). doi:10.1038/s41467-017-02261-1
118. Sarioglu AF, Aceto N, Kojic N, et al. A microfluidic device for label-free, physical capture of circulating tumor cell clusters. *Nat Methods.* 2015;12(7):685–691. doi:10.1038/nmeth.3404 [PubMed: 25984697]
119. Song Y, Huang Y-Y, Liu X, Zhang X, Ferrari M, Qin L. Point-of-care technologies for molecular diagnostics using a drop of blood. *Trends Biotechnol.* 2014;32:132–139. doi:10.1016/j.tibtech.2014.01.003 [PubMed: 24525172]
120. Williams SCP. Circulating tumor cells. *Proc Natl Acad Sci U S A.* 2013;110(13):4861. doi:10.1073/pnas.1304186110 [PubMed: 23533270]
121. Im H, Pathania D, McFarland PJ, et al. Design and clinical validation of a point-of-care device for the diagnosis of lymphoma via contrast-enhanced microholography and machine learning. *Nat Biomed Eng.* 2018;2(9):666–674. doi:10.1038/s41551-018-0265-3 [PubMed: 30555750]
122. Sundah NR, Ho NRY, Lim GS, et al. Barcoded DNA nanostructures for the multiplexed profiling of subcellular protein distribution. *Nat Biomed Eng.* 7 2019:1. doi:10.1038/s41551-019-0417-0
123. Zhang R, Le B, Xu W, et al. Magnetic “Squashing” of Circulating Tumor Cells on Plasmonic Substrates for Ultrasensitive NIR Fluorescence Detection. *Small Methods.* 2019;3(2):1800474. doi:10.1002/smt.201800474
124. Shen W, Song Y, Burklund A, et al. Combined immunomagnetic capture coupled with ultrasensitive plasmonic detection of circulating tumor cells in blood. *Biomed Microdevices.* 2018;20(4):99. doi:10.1007/s10544-018-0333-1 [PubMed: 30417219]
125. Burklund A, Saturley-Hall HK, Franchina FA, Hill JE, Zhang JXJ. Printable QR code paper microfluidic colorimetric assay for screening volatile biomarkers. *Biosens Bioelectron.* 2019;128:97–103. doi:10.1016/j.bios.2018.12.026 [PubMed: 30640126]
126. Tittl A, Leitis A, Liu M, et al. Imaging-based molecular barcoding with pixelated dielectric metasurfaces. *Science (80-).* 2018;360(6393):1105–1109. doi:10.1126/science.aas9768
127. Leitis A, Tittl A, Liu M, et al. Angle-multiplexed all-dielectric metasurfaces for broadband molecular fingerprint retrieval. *Sci Adv.* 2019;5(5):eaaw2871. doi:10.1126/sciadv.aaw2871 [PubMed: 31123705]
128. Zhang P, Wu X, Gardashova G, et al. Molecular and functional extracellular vesicle analysis using nanopatterned microchips monitors tumor progression and metastasis. *Sci Transl Med.* 2020;12(547):1–18. doi:10.1126/scitranslmed.aaz2878

129. Iyer A, Gupta K, Sharma S, et al. Integrative Analysis and Machine Learning based Characterization of Single Circulating Tumor Cells. *J Clin Med*. 2020;9(4):1206. doi:10.3390/jcm9041206
130. Sarkar S, Kang W, Jiang S, et al. Machine learning-aided quantification of antibody-based cancer immunotherapy by natural killer cells in microfluidic droplets. *Lab Chip*. 2020;2. doi:10.1039/d0lc00158a
131. Burklund A, Tadiemety A, Nie Y, Hao N, Zhang JXJ. Advances in diagnostic microfluidics. *Adv Clin Chem*. 10 2019. doi:10.1016/BS.ACC.2019.08.001
132. Grogan T, Geiss GK, Ratcliffe AL, et al. Direct multiplexed measurement of gene expression with color-coded probe pairs. *Nat Biotechnol*. 2008;26(3):317–325. doi:10.1038/nbt1385 [PubMed: 18278033]
133. Kulkarni MM. Digital Multiplexed Gene Expression Analysis Using the NanoString nCounter System In: *Current Protocols in Molecular Biology*. Vol Chapter 25 Hoboken, NJ, USA: John Wiley & Sons, Inc.; 2011:Unit25B.10. doi:10.1002/0471142727.mb25b10s94
134. Nielsen T, Wallden B, Schaper C, et al. Analytical validation of the PAM50-based Prosigna Breast Cancer Prognostic Gene Signature Assay and nCounter Analysis System using formalin-fixed paraffin-embedded breast tumor specimens. *BMC Cancer*. 2014;14:177. doi:10.1186/1471-2407-14-177 [PubMed: 24625003]
135. 510(k) SUBSTANTIAL EQUIVALENCE DETERMINATION DECISION SUMMARY ASSAY AND INSTRUMENT COMBINATION TEMPLATE. https://www.accessdata.fda.gov/cdrh_docs/reviews/K130010.pdf. Accessed November 7, 2019.
136. Biocartis. <http://us.biocartis.com/>. Accessed November 7, 2019.
137. Biocartis. <http://us.biocartis.com/assaymenu.html?elqTrackId=b5bdb9851fff4127bb822c7a932b3b2d&elqaid=10&elqat=2>. Accessed November 7, 2019.

Figure 1:

Sample-to-answer workflow within signature-based diagnostics. Patient samples are taken and analyzed with integrated microfluidic systems for biomarker isolation and detection. For quantification, advanced data processing techniques can be coupled to allow for pattern recognition and disease classification ultimately informing diagnoses.

Figure 2:

Target of key stages within the sample-to-answer workflow. Sample collection provides capture of key biological molecules such as circulating tumor cells and tumor-derived exosomes containing an amalgam of molecular information such as proteins, DNA, miRNA, and RNA. After isolation of key biomarkers, data analysis techniques can be used to elucidate presence, quantify number of molecules, or examine relative expression of multiple biomarkers towards biomarker signatures.

Figure 3:

Machine learning techniques within signature-based diagnostics. A) Shows a machine learning workflow using linear discriminant analysis (LDA) which started with training and test data for healthy, PanIN, and tumor molecular profiles (Reprinted with permission from Ref: 102). After LDA implementation, thresholds to differentiate each class was developed and a classifier developed. Classification on test data was achieved with 100% accuracy in separating the three classes. The overall sensitivity and specificity of the developed classifier was compared to a scrambled training set and one with no predictive value, showing significant predictive ability. B) Shows a similar, yet two-stage LDA classifier workflow for classifying between tumor laden and healthy individuals and then pancreatic and breast cancer patients (Reprinted with permission from Ref: 103). This approach utilized a 2-stage

threshold, in which only those first classified as tumor laden were input into the second stage where cancer subtype could be classified. C) shows a diagnostic deep network implementation which involved multicolor staining and holographic imaging for phenotyping and classification of cells (Reprinted with permission from Ref: 108). Two deep networks were utilized, the first for cell detection and the second for classification, resulting in the ability to first detect and then phenotype and classify cells from breast cancer patients. D) shows a similar two-stage deep learning integration which utilized a show a cell detection and classification deep networks to ultimately inform diagnosis of lymphoma via profiling of multiple protein surface markers within B cells (Reprinted with permission from Ref: 121).

Figure 4:

Image processing enabled machine learning techniques within signature-based diagnostics. A) shows an image processing framework built upon repeated stain/image cycles for multiplexed analysis of single extracellular vesicles (Reprinted with permission from Ref: 51). After sequential staining for multiple surface markers, merged images could be developed to profile the various tagged surface proteins. Via this profiling method, grouping of the various tested cell types could be completed using t-SNE analysis. B) shows the results of a similar workflow where multiple fluorescent markers were used to target distinct surface proteins via a sequential stain and image methodology (Reprinted with permission from Ref: 103). With this approach, relative expression of various markers within healthy individuals and breast and pancreatic cancer patients was elucidated. After introduction of

an LDA algorithm, as explained in Figure 6B, classification between healthy and tumor laden individuals, and breast and pancreatic cancer could be completed.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Figure 5:

Integrated technologies for isolation and molecular profiling of extracellular vesicles. A) shows a plasmonic nanohole assay called the nPLEX (nano-plasmonic exosome assay) which operates on the basis of extraordinary optical transmission induced by the periodic nanoholes (Reprinted with permission from Ref: 112). Upon binding of exosomes following functionalization, a shift in the resonance location is seen, and this change when normalized, can be used to determine relative expression levels of targeted surface proteins. B) shows a magneto-electrochemical sensor which utilizes magnetic beads and an 8-channel electrode configuration for the isolation and detection of exosomes. The 8-channel allows for simultaneous quantification of up to 8 different surface proteins (Reprinted with permission from Ref: 91). When these signals are normalized, expression levels of targeted surface proteins can be elucidated. C) shows a thermophoretic aptasensor which utilizes thermophoresis to aggregate fluorescent aptamer tagged exosomes (Reprinted with permission from Ref: 5). High levels of sample aggregation lead to great levels of enrichment and consequently a low level of detection. Once aggregated, based on the intensity of the resulting fluorescence, expression levels of surface proteins can be shown.

Figure 6:

Integrated technologies for isolation and molecular profiling of cells. A) Shows a triangular pillar geometry within a microfluidic device optimized for the capture of circulating tumor cell clusters by taking advantage of cell-cell junctions (Reprinted with permission from Ref: 118). The upper right figure in this panel highlights a captured cluster within the triangular pillar array with faux color and with fluorescence. With this configuration, by simply reversing flow rate, captured clusters can be eluted for downstream analysis, as shown in the bottom figure which profiled expression of transcripts. B) Shows the STAMP (DNA sequence-topology assembly for multiplexed profiling) assay and integrated microfluidic device for multiplexed sub-cellular protein targeting using barcoded DNA nanostructures (Reprinted with permission from Ref: 122). Development of barcoded DNA nanostructures allows for unique targeting of distinct proteins for highly multiplexed protein profiling and distribution analysis of sub-cellular proteins. The device and accompanying assay allowed

for molecular profiling of breast cancer patients to inform diagnoses of various molecular subtypes.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Figure 7:

Emerging technologies for multiplex diagnostics. A) Shows a compilation of related work from our lab targeting the development of diagnostic modalities that could easily be multiplexed for signature-based diagnostic applications. The top-left figure shows an Au nanorod and PNA probe optimized for detection of circulating tumor DNA point mutations (Reprinted with permission from Ref: 40). The bottom figure utilizes an Au chip and 'squashing' of CTCs via magnets to largely increased NIR fluorescence (Reprinted with permission from Ref: 123). The right figure highlights a QR-code based assay for the detection of indole, accompanied by an online cloud-based diagnostic integration (Reprinted with permission from Ref: 125). B) shows a dielectric metasurface containing an array of metaunits with varying unit cell dimensions allowing for a resonance peak map to change

over the metasurface (Reprinted with permission from Ref: 126). This surface was tuned over the mid-IR range for probing of chemical moieties. By integrating absorbance over the metasurface, barcode-like signatures unique to chemical and potentially biological molecules is achievable. C) shows another example of a dielectric metasurface which was integrated with hyperspectral imaging to allow for highly sensitive molecular detection (Reprinted with permission from Ref: 90). This study included creation of a multi-resonance sensor, which could allow for a similar barcode-like identification of key biological moieties.

Figure 8:

Diagnostic and data analysis pipelines within the development of signature-based diagnostics. The former can be sectioned into patient sample collection, biomarker isolation and biomarker detection. The latter includes advanced data processing techniques such as image processing, machine learning and computer vision. Following integration, data-driven diagnostic decisions within the clinic can be made paving the way towards next-generation personalized diagnostics.