metabolites

MDPI

Review

Lipid Profiling in Cancer Diagnosis with Hand-Held Ambient Mass Spectrometry Probes: Addressing the Late-Stage Performance Concerns

Lauren Katz [1,2,†], Alessandra Tata [3,†], Michael Woolman [1,2] and Arash Zarrine-Afsar [1,2,4,5,*]

[1] Department of Medical Biophysics, University of Toronto, 101 College Street, Toronto, ON M5G 1L7, Canada; lauren.kaufman@mail.utoronto.ca (L.K.); michael.woolman@rmp.uhn.ca (M.W.)
[2] Techna Institute for the Advancement of Technology for Health, University Health Network, 100 College Street, Toronto, ON M5G 1P5, Canada
[3] Laboratorio di Chimica Sperimentale, Istituto Zooprofilattico delle Venezie, Viale Fiume 78, 36100 Vicenza, Italy; ATata@izsvenezie.it
[4] Department of Surgery, University of Toronto, 149 College Street, Toronto, ON M5T 1P5, Canada
[5] Keenan Research Center for Biomedical Science & the Li Ka Shing Knowledge Institute, St. Michael's Hospital, 30 Bond Street, Toronto, ON M5B 1W8, Canada
* Correspondence: arash.zarrine.afsar@utoronto.ca; Tel.: +1-416-581-8473
† These authors contributed equally to this work.

Abstract: Untargeted lipid fingerprinting with hand-held ambient mass spectrometry (MS) probes without chromatographic separation has shown promise in the rapid characterization of cancers. As human cancers present significant molecular heterogeneities, careful molecular modeling and data validation strategies are required to minimize late-stage performance variations of these models across a large population. This review utilizes parallels from the pitfalls of conventional protein biomarkers in reaching bedside utility and provides recommendations for robust modeling as well as validation strategies that could enable the next logical steps in large scale assessment of the utility of ambient MS profiling for cancer diagnosis. Six recommendations are provided that range from careful initial determination of clinical added value to moving beyond just statistical associations to validate lipid involvements in disease processes mechanistically. Further guidelines for careful selection of suitable samples to capture expected and unexpected intragroup variance are provided and discussed in the context of demographic heterogeneities in the lipidome, further influenced by lifestyle factors, diet, and potential intersect with cancer lipid pathways probed in ambient mass spectrometry profiling studies.

Keywords: ambient mass spectrometry; untargeted lipidomics; untargeted metabolomics; lipid profiling; cancer diagnosis with ambient mass spectrometry

Citation: Katz, L.; Tata, A.; Woolman, M.; Zarrine-Afsar, A. Lipid Profiling in Cancer Diagnosis with Hand-Held Ambient Mass Spectrometry Probes: Addressing the Late-Stage Performance Concerns. *Metabolites* **2021**, *11*, 660. https://doi.org/10.3390/metabo11100660

Academic Editor: Guangju Zhai

Received: 23 August 2021
Accepted: 22 September 2021
Published: 28 September 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction and Problem Statement

Based on the close relationship between lipid metabolism and cancer formation/progression [1], tissue pathology determinations (cancer versus healthy or differentiation between various types of the same cancer) through lipid profiling with ambient mass spectrometry (MS) has received much traction [2–17]. A variety of aerosolization and liquid extraction methods [3] (some utilizing handheld sampling probes capable of desorbing tissue molecules under ambient conditions) are coupled to MS analysis after ionization, generating a lipid profile signature from the target tissue specimen. Through comparing the overall mass to charge (m/z) pattern of this "crude" tissue lipid profile (or signature) generated in the absence of chromatographic separation, to a library of previously collected lipid profile signatures characteristic to various tissue pathologies, rapid identification of said pathologies (cancer, infection, or inflammation) has been made possible within a few seconds of data collection and analysis [4]. The data analysis step in ambient MS often

uses multivariate decomposition methods [4] but has also been augmented with machine learning, [18] artificial intelligence [19], and deep learning methods [20] to mine further nuances in tissue classifying molecular profiles.

Targeted MS approaches for the identification of necrotic tumors using certain ceramides [21] and other unique lipids for select kinase activity monitoring [22] have been reported. While there are additional cases in which a single oncometabolite such as *N*-acetylaspartate (for differentiation of healthy neuronal tissue from glioma) or 2-hydroxyglutarate (for determination of isocitrate dehydrogenase mutation status [8,10,23–26]) is used in a targeted analysis, many ambient MS studies utilize the overall ionic pattern of tissue lipids in an untargeted analysis or a subset thereof (e.g., most distinguishing ions as from a parsimonious method that uses feature reduction [27]) as the distinguishing "biomarker" [2–5]. These developments implicitly redefine the "classic" notion of a disease biomarker from a single (often protein) molecule that is reproducibly validated across many specimens in a "targeted" analysis, to a pattern of disease specific (and thus disease classifying) metabolites or lipids analyzed in their unique ionic states in an "untargeted" manner, often acquired across far fewer independent specimens. The new "biomarker", in other words, is a multidimensional classifier that is represented by its coordinates on a multivariate decomposition plot. The act of disease identification using said classifier thus boils down to tracking the juxtaposition of the multivariate coordinates of a query mass spectrum on the multivariate decomposition plot of the reference signature library. Here, a pre-defined certainty level to call out positive hits (overlapping with the model) from non-overlapping (or negative) hits is used.

To date, diverse ambient MS methods have been employed to address a wide variety of clinical cases [2–5]. Overall, very good analytical performance metrics for lipid or metabolite pattern matching in ex vivo or in situ tissue explorations across tens to hundreds of independent clinical specimens have been reported [2–5]. Often, by simply splitting the acquired (albeit highly limited) datasets into training (model building) and test (recognition) subsets, predictive multivariate classification models are created and validated based on statistical association with the disease state signal. The success of ambient MS profiling in the context of such limited analytic validation steps has largely been hailed as resulting in promising new avenues for rapid and accurate cancer diagnosis. Beyond lipidomic analysis of solid tissues with untargeted ambient MS methods described above, utility of plasma lipidomics [28] for cancer diagnosis is also gaining momentum [29]. A detailed review of technologies and various applications of ambient MS sources (beyond the scope of this manuscript) is provided elsewhere [2–5]. However, in summary, an ambient MS analysis source often employs an extraction mechanism, desorption−ionization, or a desorption followed by post ionization wherein the specimens are subjected to extraction of molecular content largely in the absence of extensive pre-processing and under ambient conditions (Figure 1A). A variety of extraction or desorption methods are used, leading to a diverse set of applications ranging from clinical diagnosis to food safety and pharmaceutical explorations (Figure 1B). Laser-based ambient ion sources have been extensively reviewed recently by our group [30], and need not be repeated here. It must, however, be emphasized that as of late, the majority of hand-held ambient MS probes (focus of this manuscript) do not utilize pre-separation of analytes (e.g., by means of chromatographic or ion mobility). As such, direct infusion (as done in conventional lipidomics) best describes the state of the art in ambient MS methods. This poses a fundamental challenge in the applicability of analysis tools such as Compound ID and content look up databases such as LipidMaps to ambient MS data.

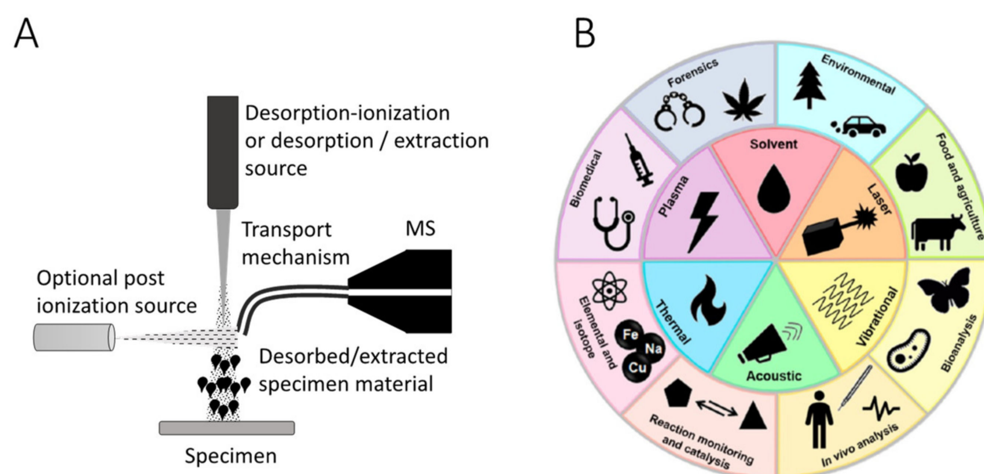

Figure 1. Principles and high-level overview of the mechanisms of ambient mass spectrometry. (**A**) Schematics of a typical ambient MS source. A desorption−ionization source or an extraction source provides desorbed or desorbed and ionized molecules from a specimen under ambient conditions. Inclusion of post ionization is optional and enhances the signal in select cases. A mechanism for the transport of extracted/desorbed materials to the mass analyzer is included that may involve flexible tubing for long (>2 m) transport of surgical aerosols as in iKnife [14], laser desorption plume as in SpiderMass [31] and PIRL-MS [32], or water (as solvent) extracted tissue content as in the MasSpec Pen [17]. The ionization step can take place anywhere between the specimen surface or close to the mass analyzer. A notable difference is lack of chromatographic separation of analytes prior to MS analysis. (**B**) A high-level overview of core ambient desorption/ionization technologies and related applications. Here, a variety of desorption/ionization methods consistent with the schematics provided in panel A are used to generate a multiple application base that ranged from surgical to material to environmental or forensic explorations. Reprinted (adapted) with permission from [3], Copyright 2019 American Chemical Society.

By the same token that a single disease biomarker molecule must pass rigorous analytical, regulatory, and late-stage performance surveillance (population validation) to ensure high sensitivity and specificity, lipid and metabolic profiling with ambient MS must also be subjected to the scrutiny of the same guidelines and principles. In this quest, lessons can be learned from the failure of many protein biomarkers to reach the clinic [33–35], and these lessons can be applied to the nascent field of rapid tissue pathology through lipid profiling with ambient MS. Five important differences set lipid profiling with hand-held ambient MS probes apart from many single molecule (protein) biomarker discovery and validation efforts. First, the majority of published ambient MS studies used a small, and rather homogenous population of clinical specimens in the analytic validation step to assess the accuracy, specificity, and sensitivity of the tissue pathology predictions [36]. While data analysis methods have been introduced to boost predictive power of limited models using “cumulative learning” approaches with some success [37], the issue of a model’s poor predictive ability, arising from small size, deserves special attention. Second, tissue lipid profiles are likely to be influenced by intrinsic population genetic heterogeneity, diet [38–40], and other tumor microenvironment factors such as stromal content and presence of hypoxia, among others (Figure 2). It has been shown that changes in diet and lifestyle interventions alter tissue lipidomes in mice, but not all tissues are affected to the same extent or show the same altered lipidomic pattern [41]. Third, our collective understanding of the underlying lipid metabolic pathway heterogeneities [42] remain limited, as few studies have investigated large-scale variation of lipidome across healthy populations. Fourth, most published ambient MS reports fail to further determine the molecular identities of important tissue classifying lipid molecules. This makes utilization of metabolic pathway knowledge in a rational approach to further refine them to those least influenced by population heterogeneity and diet a daunting task. Fifth, by definition, tissue

lipid profiling is an "untargeted" MS approach that, in the absence of chromatographic or ion mobility separation, is vulnerable to population noise perhaps to a greater extent than a single known and well characterized molecule in a targeted approach would be. While parsimonious analysis methods that use feature reduction may be less susceptible to population noise compared with wide mass range models, discovering a mechanistic link between biomarker and disease state further buttresses biomarker validation beyond just what is inferred from the strength of its statistical association with disease state. Statistical association as the sole criteria for biomarker identification and validation in an untargeted approach can be misleading if not supported by additional mechanistic verification [35]. The unequivocal identification of metabolites must utilize high resolution mass spectrometry, tandem MS analysis for diagnostic fragment identification in conjunction with additional separation (ion mobility and/or chromatographic) methods as an orthogonal approach.

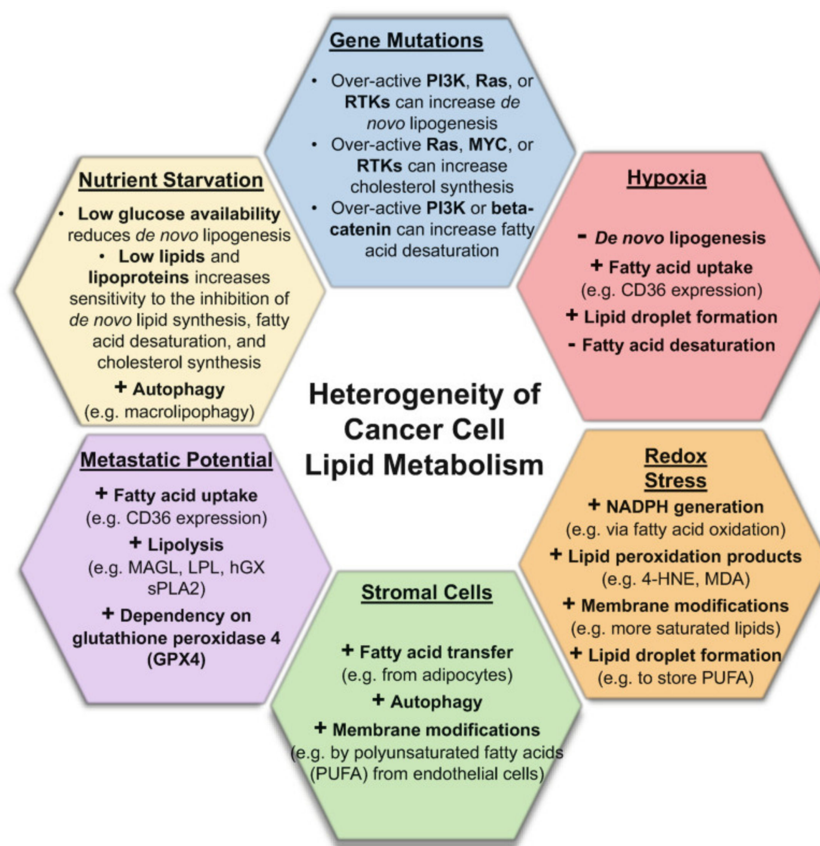

Figure 2. Heterogeneity of lipid metabolism in cancer. Overview of multiple factors that can influence lipid metabolism in cells, creating additional molecular heterogeneities that must be captured in ambient MS profiling studies through use of a large, balanced, and heterogenous sample set. These factors can be external (diet) and can arise due to intrinsic biological architecture (stroma), mutations, and the tumor microenvironment (hypoxia). Reproduced with permission from [43].

This review provides a critical assessment of the dangers associated with the use of small initial (and often unbalanced) sample sets [44–47] in pattern recognition [36] and lack of late stage validation with respect to the utility of the discovered lipid pattern for a high sensitivity and specificity performance across a diverse population [36]. We then accordingly provide recommendations drawn from the growing field of untargeted MS analysis in food sciences facing similar challenges in validation strategies [48–51]. To detect an unlimited set of potential adulterants with an untargeted approach, a large and diverse population of authentic products must be examined [52]. We first review broad stroke principles of analytical, late stage, and regulatory validations, and then attempt to extend the implementation of predictive markers from early phase trials [53] to large scale

validation [54], drawing further upon sources of failures in the protein biomarker discovery world [33–35,55,56] to provide validation strategies for ambient MS profiling that are most suitable to minimize late stage performance concerns across diverse populations.

2. State of the Art

Any method intended for clinical decision making or use at the bedside must offer, over a pre-defined intended use, a precise, accurate, robust, sensitive, and specific performance within a reportable range that is reproducible across many sites, users, specimens, and is further uninfluenced by instrumental and environmental factors and duty cycle (carry over), and should additionally meet specific safety and use site compliance requirements [57,58]. These constitute analytic performance matrices that have recently been reviewed in the context of ambient MS method use at the bedside [57]. When laboratory developed tests (LDTs) [59] are used to develop a new targeted MS approach, often, a homogenous, relatively small sample set sourced from patients local to the laboratory site or region may be used as the first step for validation. This strategy, however, leaves the analytic validation step vulnerable to population variation, leading to late-stage performance concerns not addressed during first trials. While the pharmaceutical industry rigorously engages in post-marketing surveillance campaigns to monitor the performance of novel therapeutics (in terms of safety and also efficacy), LDTs, if not validated using a sufficiently diverse (or statistically reduced or demographically matched) test case, remain very vulnerable to population level poor performance [60], as seen in a number of failed biomarker cases that showed good performance in the laboratory but failed to perform well across the population in the absence of demographic matching [54–56]. Here, limited information such as changes in a marker in response to disease progression, pharmacological intervention, and crosstalk therein, will further complicate the co-efficient of similarity-based methods in the absence of careful demographic matching with the model. While biomarker driven targeted approaches (or those with statistically reduced datasets) that rely on identified disease specific peaks may be somewhat shielded against such variations, similarity coefficient-based comparison methods that rank a query spectrum against the spectra from a collection of diseases must be more closely scrutinized. Currently, two untargeted MS methods for pathogen strain identification have broken through the regulatory barrier [61,62]. Building on a possible pathway of LDTs [59], the USP Pharmacopeial Convention provides guidelines for analytic validation of untargeted mass spectrometry methods in food sciences, quality control, and for the detection of adulteration [49–51,63]. Further resources for bioanalytical method validation are available through the FDA [64], and a recent adaptation of pharmacopeial and bioanalytical methods validation strategies specifically tailored to ambient MS methods has been recently published [57] in an attempt to standardize an analytic validation workflow for this technique in cancer research. This review (by our group) attempted to put forward guidelines for rigorous analytic validation, including suitable sample sizes [65] according to power calculation [66], as recommended for metabolic phenotyping [44,45] and stayed faithful to the minimum reporting standards for chemical analysis [46,67]. However, late-stage validation concerns for rapid tissue lipid profiling that use lessons learned from failures of targeted single molecule protein biomarkers to reach the clinic are not reviewed in depth, and thus form the basis of this report. Moreover, a previous report by our group qualitatively addressed the influence of tissue molecular heterogeneity on ambient MS profiles in the context of choosing the correct disease model [68]. This review further details the origins of this heterogeneity and provides recommendations for optimal design of validation studies to shield against the potential confounding effect of population heterogeneity on ambient MS profiles.

2.1. Molecular Heterogeneities in Biological Tissues Impact Metabolome and Lipidome Profiles

As discussed in the section above, lipids have been shown to play diverse roles in many cellular functions. It was postulated (a decade ago) that lipidomics would evolve to play a critical role in expanding our understanding of disease states [69]. As speculated then, lipidomic technologies have indeed matured considerably to broaden our understanding of how lipid metabolism correlates with disease biology [70], especially in cancer, where lipids [71,72], in addition to being of diagnostic value [29], may be involved in a number of regulatory pathways [43,73], thus showing additional promise as therapeutic targets [74,75]. While our ability to effectively target cancer through modulating lipid metabolism is still an open question [76], there is convincing evidence that there exists a significant amount of molecular heterogeneity in cancer cells [77,78]. The exact ramification of this molecular heterogeneity in terms of divergent metabolic (or lipidomic) pathways is only beginning to emerge in select cases [43,78–87], but there is convincing evidence that many factors, including adaptive evolution (to treatment [88]), microenvironment components (stroma and inflammatory cells), and hypoxia, among others, can influence cancer cell metabolism [89], and in some cases directly affect lipidomic profiles [43,81,84,90–92]. Of notable importance is a study of lipid metabolism in breast cancer that suggests a differential role of essential and non-essential lipids in metabolic profiles of inflammatory and cancer cells within the tumour tissue [91]. In the absence of spatially resolved ambient MS analysis (or single cell lipidomics [93]), to better capture these heterogeneities, alongside those arising from metabolic adaptability of cells to their environment [94], at the very least, a large population base must be included in the analysis, particularly in the initial model building stage. The issue of spatial resolution (often not at single cell levels in prominent ambient MS methods) causes sensitivity of signal to tumour cellularity, as shown in a DESI-MS lipidomic profiling study of brain cancers [95] and the localization of high 2-hydroxyglutarate (2HG) signal levels in the regions of glioblastoma tumours with a dense cellularity [10]. While this sensitivity can be utilized to estimate tumour cell percentage [96], which is in its own right a useful clinical indicator currently only obtainable by time consuming histologic staining and microscopy, caution must be exercised in the interpretation of ambient MS profiles associated with residual disease presence. Residual HER2 expression within breast tumours [97] in a proteomic MALDI-MS study has been shown to influence receptor status predictions [98].

In a recent study aimed at creating predictive molecular models of breast cancer receptor status with ambient DESI-MS [99], intrinsic molecular heterogeneity in HER2 receptor patients [100] resulted in dramatically reduced accuracy of DESI-MS models for the determination of this receptor status compared to that of the estrogen (ER) and progesterone (PR) receptors from DESI-MS models. In a similar vein, in a previous ambient MS profiling study of pediatric medulloblastoma cancer molecular subgroups, an aberrantly high content of stroma or an infiltration of healthy tissue in the specimens to be classified, not accounted for in the model, led to misclassification or failed classification [9]. This study highlights the significance of comprehensive molecular models that capture the expected molecular heterogeneity associated with the infiltration and microenvironment. Further emphasizing the impact of the architectural (and hence molecular) heterogeneity of cancer, DESI-MS modeling of ovarian high-grade serous carcinoma and serous borderline ovarian tumours resulted in misclassification of a few specimens that contained microcapillary growth patterns and other architectural complexities common to invasive carcinoma features associated with a specimens' unique pathology [101].

In addition to these discoveries that highlight the extent of the influence of tissue molecular heterogeneity on the predictive power of MS-based modeling, it has long been recognized that diet may modulate both plasma and certain tumour lipids [102]. Diet exchange between demographic groups has been shown to alter the metabolome (and cancer risk [103]), and chronic exposure may mark an effect on the metabolome [104]. Cancer cells are known to scavenge nutrients from their environment, especially under aggressive growth or inadequate perfusion [105], creating an additional link between diet

and metabolism. Furthermore, the availability of precursor lipids diffused to the tumour site [106] is shown to create additional spatial heterogeneities in tissue metabolic states in the vicinity of cancer [107]. A high fat diet has been shown to alter lipid metabolism in cancer and influence the interplay between normal and cancer cells [108]. This relationship is exemplified elegantly by the crosstalk between adipose tissue (storage for dietary lipids [109]) and tumors [110], with cancer-associated adipocytes [111] providing adipose-derived lipids for cancer progression [112] and proliferation [113], with stipulated therapeutic potentials [114,115].

Further complicating the interplay of diet, adipocytes, and lipid uptake in cancer metabolism discussed above, alcohol consumption is shown to perturb adipocytes, promoting lipolysis [116–118] and thus potentially influencing cancer cell metabolism by modulating released precursor lipids. In addition, there has been a further recognition of the influence of gender and age on the composition of the plasma lipidome [119–121]. While the influence of plasma lipidome variations on solid tumour profiles is yet to be systematically investigated, a picture thus emerges that to improve the diagnostic accuracy of lipid profiling for cancer characterization with ambient mass spectrometry, diet, gender, and lifestyle factors, among others (e.g., clinical history regarding treatments received) must be taken into consideration, and that ideally, spatially resolved analysis is required to capture intratumoural lipidomic or metabolomic heterogeneities. While the latter is not possible or technically feasible, significantly large sample numbers to capture as much of the population level or batch effect variances [57] will be helpful.

2.2. Initial Statistical Modeling Should Be Based on Sufficient Sample Numbers

The impact on lipidomic profiles of tissue molecular heterogeneity discussed above further supports the much needed emphasis on suitable sample size to produce significant results in pattern recognition [36,122,123] and related efforts [36], as well as in multivariate and high dimensionality data analysis [36,44,45,47]. Effective sample numbers that result in statistically significant profile models must be optimized for each case study and incorporated in the earliest phase of ambient MS profile model building. Here, inter- and intra-group variance in lipid profiling data is a key determinant of the overall sample sizes required to differentiate between closely related or molecularly distinct tissue types, especially if elevated levels of intra-group heterogeneity and variance are present in the dataset. In Figure 3, we illustrate this point by re-analyzing published data [9] corresponding to the principal component analysis and linear discriminant analysis (PCA-LDA) of 4-component groups, closely related and molecularly distinct. As illustrated in this figure, larger sample numbers per group are required to effectively distinguish closely related groups compared to molecularly distinct groups. Here, analysis of so-called "learning curves" in the related domain of micro-array analysis [124,125] have shown promise in determining effective sample sizes [126], and a minimum of 75 specimens per class have been suggested [127]. Here, there are parallels to draw from proteomics using Matrix Assisted Laser Desorption Ionization Mass Spectrometry (MALDI-MS) profiling [128–131], where guidelines for determining the appropriate sample size for clinical proteomic profiling studies have been put forward, using a linear mixed model that allows for the inclusion of estimates of biological and technical replicate variance in a given experiment [132]. In a similar vein, by extending previously published approaches [133,134], the effect of biological variance (e.g., intratumoural heterogeneity), as an important determinant of the overall sample size required, has been examined towards a method that allows for the adjustment of expected biological variance in calculations of sample size [135]. A central tenet of this approach is a rigorous replicate analysis (i.e., multiple sampling) of each specimen towards determination of the intraclass heterogeneity. Here, averaging replicates and using this value in class differentiations is discouraged as averaging results in loss of valuable information regarding intrasample variability [135]. More interestingly, beyond basic multivariate methods, such as linear discriminant analysis (LDA) [136], least absolute shrinkage and selection operator (LASSO) [137], machine learning approaches such as

support vector machine (SVM) [138], and random forest (RF) [139], which require extensive pre-processing of mass spectral data, convolutional neuronal networks have been proposed to offer a higher accuracy of prediction without the need for data pre-processing [140,141]. The predictive power of these methods, however, sharply decreases with small sample sizes [142,143]. While cumulative learning with convolutional neural networks have been proposed to utilize smaller mass spectrometry datasets [37]; the future augmentation of ambient MS data analysis methods with artificial intelligence is bound to make a larger demand on diversity of training datasets and careful considerations of appropriate sample sizes as discussed above.

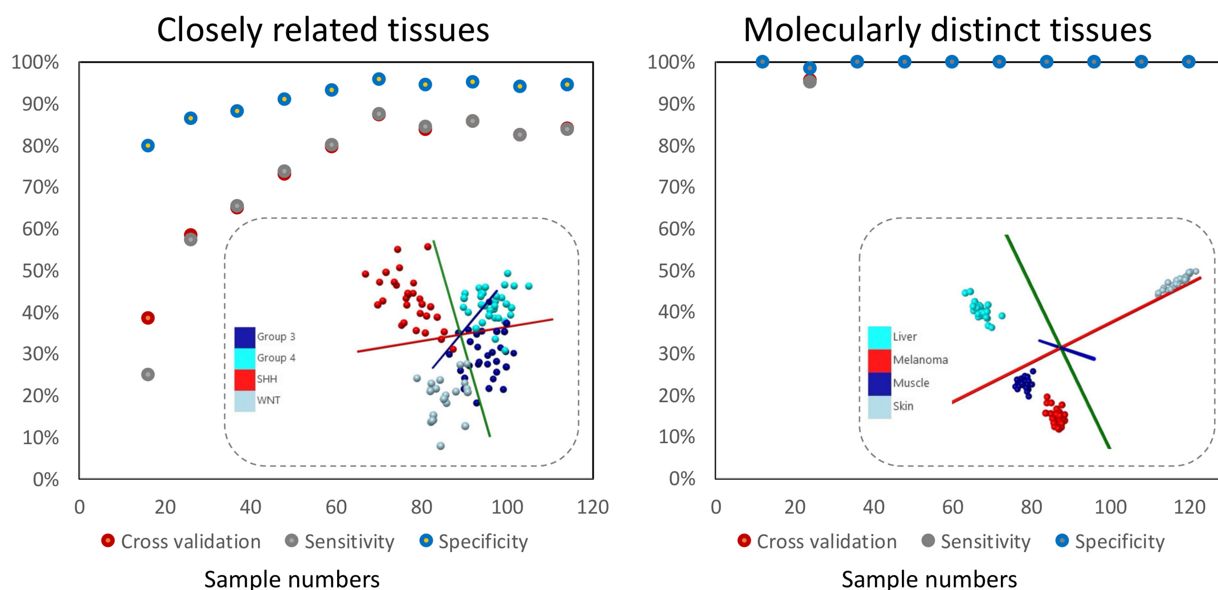

Figure 3. The influence of sample size on the predictive power of ambient MS methods. In this figure we re-analyzed previously published [9] and newly acquired picosecond infrared laser mass spectrometry (PIRL-MS) data from two cases of 4 closely related tissues (subgroups of human pediatric brain cancer medulloblastoma) and 4 distinct tissue types (murine melanoma and additional liver, muscle and skin tissues). Each dataset is comprised of a 10-s ambient PIRL-MS sampling (m/z range 100–1000 Da) subjected to principal component analysis and linear discriminant analysis (PCA-LDA), as described previously [9]. Cross validation was performed on AMX [144] using a 20% leave out [9], and was used to calculate the specificity and sensitivity. For each specimen, one 10-s PIRL-MS spectrum has been included. However, in keeping with the recommendations of Nyangoma et al. [135] (see above), multiple sampling of each specimen towards developing the extent of intrasample variation has been reported in our original publication [9]. As can be seen here, where there is significant intrasample variability (large spread of data points for each class in the PCA-LDA scores plot) and where classes are more molecularly alike, larger sample numbers are required to reach prediction power plateau compared with cases where less variant, molecularly distinct classes are compared. In this comparison, each class contained ~30 specimens (and one PIRL-MS spectrum per specimen).

2.3. Lessons Learned from Metabolomic and Proteomic Biomarker Discovery and Food Sciences

The horizontal growth of untargeted ambient mass spectrometry profiling has taken place in the absence of an effective dialogue with researchers in the areas of metabolomics or proteomic biomarker discovery. While most of the "conventional" biomarker discovery studies use targeted approaches, there are challenges in metabolic profiling study design (reported previously [145,146]) that may be applicable to the untargeted ambient MS analysis of tissue molecular profiles. It must be emphasized that a large fraction of true biomarker discoveries (those that have been validated in controlled laboratory or through early-stage discovery or small population studies) fail to break through the clinic [33,35,55]. While the sources of such failures can include many factors such as low added value to prognostic or clinical utility (stemming from improperly defining the clinical need) [33], underestimating cancer heterogeneity (extreme case selection) among others reviewed elsewhere [55,56], a

low sensitivity or specificity across the wider population [54] in late-stage or multisite trials have also been reported [55]. As an example, B7-H4, a novel membrane bound protein proposed as a marker for ovarian cancer [147], failed due to the large variance seen across multisite validation efforts [148]. Further buttressing the importance of careful study design, lysophosphatidic acid (LPA) [149] failed to become a reliable marker for gynecological cancers as the debut publication used a non-standard sample processing protocol [150]; thiosulfate failed to stand the scrutiny of additional benchmarking against other orthogonally validated markers for prostate cancer despite early stage promise [151]. Ambient MS profiling is not immune to the hurdles and challenges reviewed elsewhere [33,55,56], especially those that highlight the importance of large study validation using a diverse population [54,55], reducing the hype based on limited initial findings [55]. Extrapolating parallel sources of failures seen in these fields [54–56] to the design of ambient MS profiling studies will be helpful. Table 1 summarizes a few key points of value for consideration by investigators using ambient MS profiling in their research. In essence, biomarker discovery efforts may face pitfalls in all stages of discovery, validation, translation, evaluation, and implementation. As illustrated in Table 1, small studies or those that use extreme cases without careful validation and use of replicates, rationalization of discordant results, or careful design of a randomized trial across sufficiently diverse set of parameters are likely to fail. In a similar vein, a poorly defined clinical case that lacks clear added value to the clinical practice is unlikely to pass the evaluation or implementation stage. Lastly, solutions that pass the scrutiny of the above are not necessarily bound to succeed either. A solution that passes implementation must improve outcomes or offer performance added value across diverse settings and must also not be cumbersome to implement (be easy to use). Here it must be emphasized that properly defining the key parameters listed in Table 1, such as "diverse setting", "sufficiently large datasets", "ease of use", and "clear added value" is not an easy undertaking. Clearly defining these factors requires an intimate understanding of the underlying diversity in the molecular make up of target molecule(s) across a large number of factors, a feat that is far more complicated in untargeted analyses. Here, the untargeted metabolomic analysis of adulterated food provides helpful parallels, especially in revealing the importance of capturing heterogeneity in the initial statistical model studies. Due to the untargeted nature of the validation protocols, either a vast number of potential adulterants must be included in the initial model building efforts, or a large variety of authentic products from diverse origins must be studied and their molecular fingerprints catalogued as authentic [49]. In food authentication, the US Pharmacopeia guidelines explicitly note that samples in the reference set must fully encompass all of the variability of the food product under study [52]. Large sets of samples enable the appropriate inclusion of biological variance, and further including replicates is encouraged to minimize unexpected variations [49,51]. This scenario closely parallels lipid profiling studies of highly heterogenous cancers wherein as much of the existing inter- and intra-specimen heterogeneities as possible must be captured. Further emphasizing the importance of comprehensive reference datasets that capture the entire range of expected specimen heterogeneity, in an evaluation of certified reference materials used for oregano authentication with untargeted ambient mass spectrometry, a false positive was identified; the model failed to classify a certified sample from South America as authentic due to the training set utilized having been only populated with certified oregano specimens with a European origin [152]. Similarly, two newly established untargeted molecular models for extra virgin olive oil authentication with gas-chromatography fingerprinting failed to classify aged samples from a previous harvest season [153]. The limitation of insufficient heterogeneity in the reference dataset was overcome by focusing on the presence of adulterant sunflower oil instead, in order to assess olive oil samples from new geographical regions not included in the reference model [154]. This strategy was suitable as there was only one type of known adulterant expected, and may not be applicable to all ambient MS profiling studies. Here, a comprehensive review of the challenges associated with untargeted fingerprinting methods used in food authentication has been published that

emphasizes frequently updating the model and revising them as new data become available [155]. This puts forward a new vision for the creation of intelligent, self-improving molecular models for cancer diagnosis with untargeted ambient MS lipid profiling wherein the availability of new data post analytic phase (implementation) can result in naturally evolving models that capture additional heterogeneity from diverse sites, and across varied socio-demographic strata. A key to the successful implementation of this vision, besides artificial intelligence models, is a rigorous definition of ground truth (i.e., pathology) to ensure only spectra from validated authentic samples are included in the revised dataset.

Table 1. Challenges and sources of potential failures in conventional targeted biomarker discovery efforts. This table lists current challenges in different phases of biomarker discovery. Parallels, at each phase of discovery, validation, translation, evaluation, and implementation can be drawn to be used as guidelines for potential pitfalls of ambient mass spectrometry profiling. Reproduced with permission [55].

	Current Problems	**Potential Solutions**
Discovery	Poor design, conduct, and analysis	Methodological rigor
	Unaccounted multiplicity	Appropriate use of statistics
	Small studies	Larger, collaborative studies
	Extreme case selection	Proper case-control or cohort selection
	Nonrigorous exploratory nature of studies	More rigorous training of scientists
	Poor reporting	Use of reporting standards
	Selective reporting	Preregistration
	Spin in interpretation	Careful editorial and peer-review
Validation	Any and all problems seen in discovery studies	Similar solutions, as above
	Lack of replication efforts	Incentives for running replication studies
	Inbred replications (same populations, same investigators)	More emphasis on external, independent validation
	Incomplete, suboptimal validation	Careful consideration of independence
	No systematic reviews	Good-quality systematic reviews
	Inflation in early, small studies	Large validation studies, ideally from collaborations without bias
	Spurious variability in measurements, methods, analyses across studies	Standardization and harmonization of processes, collaborative consortia
Transition to clinical translation	Inappropriate perusal of clinical translation	Rigorous systematic reviews
	Poor prioritization	Rigorous umbrella reviews
	Sponsor bias driving translation urge	Independent assessment of the evidence
	Inappropriate stagnation without clinical translation	Incentives to translate

Table 1. *Cont.*

	Current Problems	Potential Solutions
Evaluation	Focus only on accuracy and process measures	Emphasize patient outcomes
	Few randomized trials of biomarkers	Promote randomized trials of biomarker use
	Use for unclear informational purposes	Evaluate utility of information for the sake of information and potential collateral harms
	Improper use for selection and stratified/subgroup analyses in trials	Validation of utility of stratified/subgroup analyses
Implementation and deimplementation	Poor understanding of the use of biomarkers in real-life settings	Implementation studies assessing use and outcomes in diverse settings
	Lack of rigorous guidelines	Standardized, nonconflicted guidelines
	Discordant guidelines	Strengthening of regulation for biomarkers
	Not well-defined regulatory landscape	Testing of utility of long-used biomarkers
	Entrenched useless biomarkers	Overcoming resistance from conflicted stakeholders, higher barrier for reimbursement
	Resistance to deimplementation even with convincing negative evidence	

2.4. Moving beyond Statistical Associations

A caveat of supervised untargeted ambient MS analysis is its strong dependence on statistical correlation of mass spectra with ground truth information (e.g., pathology). The untargeted nature of lipid profiling that often utilizes many features (m/z values) creates additional vulnerability to population level variations. While restricting the mass range to smaller regions may provide some benefit, the analytic validation guidelines summarized above (and discussed previously [57]) do not per se call for identification of tissue classifying lipids. Therefore, establishing a further mechanistic link between lipid profiles and disease state is not always performed. When small sample sizes pose a challenge on study significance [36], boosting confidence in association of lipidomic patterns with the disease state uncovered in the validation phase through a mechanistic link will be advantageous [156]. Furthermore, as our knowledge of the metabolic pathways [42] advances [157,158], controlled xenograft models continue to provide an additional test-bed for orthogonal validation of certain lipids involved in metabolic pathways [159–163], containing selective mutations in lipid synthesis machinery [22]. Feature reduction to perform disease classifications with as few tissue-specific and highly distinguishing lipids as possible (through low complexity or sparse analyses) [9] is an attractive strategy to further shield against population noise that may alter the spectra without influencing all the strongly classifying peaks. This rational approach further allows for the inspection of potential crosstalk between dietary pathways (food metabolome [164]), and identified lipid marker(s) pathways to anticipate heterogeneities (utilizing human metabolome database insights [165–167]) or to iteratively reduce m/z features towards development of an initial model that is robust in terms of tissue classification and uninfluenced by population level noise or heterogeneity. An example of this is an investigation of molecular signatures of ischemic heart tissue (myocardial infraction) using spatially resolved ambient MS profiling with machine learning mining and molecular identity determinations, wherein a role for taurines in the infarction process was discovered [168], consistent with the physiological role of taurines in myocardial tissues [169]. A growing body of evidence suggests that diet-induced depletion of taurines influences cardiomyopathy and its supplementation is

helpful in attenuating the degradation noted [170,171]. Rational investigation of pathways in the context of molecular identification could thus prove beneficial in revealing potential crosstalk with diet. Here, an understanding of the basal level variation across healthy population will be an important step in rationally creating suitable molecular models that capture a sufficient level of heterogeneity for the purpose of accurate predictions.

In addition to providing further support for involvement of a particular metabolite (detected with ambient MS) in a disease process, metabolic pathway information can also be utilized to enhance classification by invoking and subsequently involving additional metabolites, rationally determined to be influenced by said disease pathway, in the classification process. An example of this is reported in a recent work from the Zare group, where rational inclusion of cognate Krebs cycle metabolites, also involved in other cancers [172], in the classification algorithm not only validated the initial set of discovered metabolites relevant for basal cell carcinoma detection, but also improved the classification accuracy of the model used [137].

3. Recommendations and the Proposed Workflow for Ambient MS Method Validation for Rapid Pathology Determination

Based on the points discussed above, and taking an exemplary application of rapid pathology readout, in Figure 4, we propose a workflow for ambient MS analysis of biological tissues. Some of the elements in the proposed workflow may have to be further defined in conjunction with notes of caution and solutions provided in Table 1. We nevertheless propose to, in a general and high-level sense: (1) Define the clinical need (or added value) by performing thorough key opinion leader interviews at the earliest stage possible. For example, non-subjective tissue pathology information from ambient MS is useful. However, whether current standard of care methods have any shortcomings in terms of speed, accuracy, or scope of available information to be delivered by ambient MS must be established first. Beware of the fact that clinical decision making in the standard of care workflow is based on the scope of information currently obtainable. While rewriting the standard of care around new information (now available) is possible, strong justifications (and possibly parallel outcome data) are required to suggest revisions to the standard of care workflows. Rapid diagnosis will not always drastically improve clinical decision making. (2) Use a diverse and balanced sample set at the first possible (initial) modelling effort opportunity. This will allow for the sensitivity and specificity (predictive value) to be established using a heterogenous set early on. Verify how predictive power changes as a function of sample numbers included in the model and increase the sample numbers until the predictive power has reached a plateau (in a "predictive power" versus "sample number"correlation plot). If possible, establish population noise level by including a cohort of healthy specimens in the study and pay special attention to intraspecimen heterogeneity by performing multiple samplings across each specimen. Likewise, perform blind test validation with a diverse set, ideally sourced from a different demographic origin (see Table 1 for pitfalls and redocumentations). (3) Perform a low complexity (or sparse) analyses, and feature reduction to define the minimum number of lipids that can perform the classification without drastically sacrificing predictive power, sensitivity, and specificity. (4) Perform identification of most distinguishing (or tissue-classifying) lipids using orthogonal tandem MS or high resolution, chromatography-enabled MS analysis in conjunction with tools and platforms created for conventional lipidomics or metabolomics such as LipidMaps [173]. HPLC-MS/MS analysis can utilize "pathology guided" tissue section sampling methods such as those with laser capture microdissection. It must be noted that a rational association between markers and dysregulated disease pathways will further enhance feature reduction. Here, correct annotation of identified lipid markers is of exceptional importance also for augmenting the current knowledge of disease pathways and their functional decoding. Unfortunately, inaccurate annotation is common in many lipidomics studies [173,174]. This has created a need for standardization [175] where the use of latest guidelines [173] is recommended. (5) Create a mechanistic link between identified lipids and disease biology if feasible (and validate using engineered xenograft models) and inspect metabolic pathways for possible

crosstalk with dietary lipids, performing additional feature reductions towards a set of lipids that provide robust predictive value, specificity, and sensitivity with less expected crosstalk with diet. (6) Evaluate the predictive model of full spectrum and feature reduced models using the validation set and use the model that provides best performance and use this set for further multi-site, multi-user validation by exchanging standardizing and following devised protocols interpretable by multiple users at different sites. Lastly, we highly encourage investigators to define their own user skill set requirements and establish early on who will be implementing the solution; do they possess the right skillsets or find the solution difficult to implement? As stated above, a combination of Table 1 notes of caution and the high-level summary provided in Figure 4 must be considered to address the late-stage performance discordance in adopting an untargeted ambient MS method for bedside use.

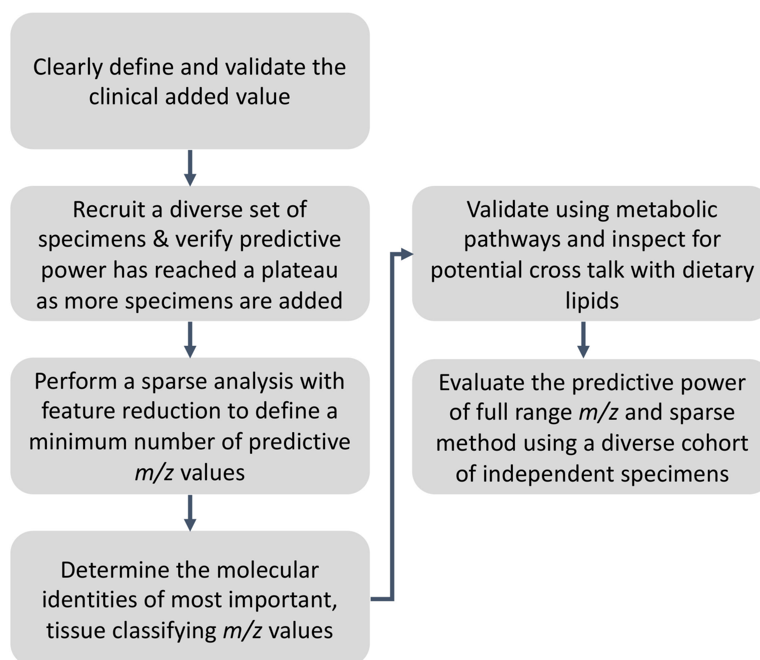

Figure 4. Recommendations for ambient MS profiling workflow focused on utility for rapid pathology determination. Summarizing the material gathered and discussed in this review and major sources of failures in conventional biomarker discovery efforts (Table 1), we have made a high level workflow for ambient MS analysis. This emphasizes proper definition of a useful clinical case, use of diverse sample set that takes into consideration intraclass heterogeneity using rational statistical sample size calculation methods discussed in the text, alongside an impetus to go beyond statistical associations to identify tissue classifying molecules and rationally justify their association using common disease pathway knowledge. This has the added bonus of feature reduction to reduce crosstalk with obvious metabolic heterogeneity factors such as diet. These recommendations are only high-level and should be used in conjunction with lessons learned from Table 1, and material detailed in the text, especially with respect to recommended sample size calculation methods. Providing a one-size-fits-all recommendation for a suitable sample size calculator is beyond the scope of this paper; sample sizes must be optimized for each study.

4. Conclusions and Caveats

This review paper aims to encourage a dialogue between investigators that use MS profiling for cancer detection and those involved in clinical biomarker validation on the population level. Using a limited initial sample set and validating it with a homogenous specimen set that likely fails to capture the large-scale variance expected at the population level constitutes an important misstep that has led to the failure of many protein biomarkers to reach the clinic. We provided six recommendations to encourage robust and early careful

study design that takes into considerations pitfalls of promising protein biomarkers that failed to reach utility at the bedside, to facilitate clinical translation of rapid lipid profiling for accelerated cancer diagnosis in the clinical domain. Poor definition of the clinical added value, emphasis on statistical association in the absence of careful statistical calculations to justify sample sizes needed among others, alongside poor understanding of the influence of lifestyle and diet on cancer lipid profiles constitute important areas of improvement in future data modeling. We hope that our recommendations will save time and effort in the early evaluation of promising leads towards successful translation in a manner that reduces the discordance between initial phase promise and late-stage performance. It must be emphasized that the recommendations provided in this manuscript aim to enhance untargeted ambient MS methods that do not utilize pre-separation of analytes (as done in conventional lipidomics), such as those obtained with hand-held MS probes. As of late, ion mobility separation is being incorporated into many MS sources, which will undoubtedly enhance the robustness of ambient MS datasets. Therefore, appropriate revisions to these guidelines are envisioned as the field of ambient MS evolves to incorporate new separation technologies.

Funding: This research was funded by the Natural Sciences and Engineering Research Council of Canada (Discovery Grant Program, RGPIN-2018-04611).

Data Availability Statement: The data presented in this study (Figure 3A, re-analyzed previously published results [9]) are available on request from the corresponding author. The raw data is not publicly available due to Ethical research guidelines that prohibit sharing and transfer of anonymized human specimen raw data without institutional authorization.

Acknowledgments: Research in mass spectrometry in the Zarrine-Afsar group is currently (2021) supported by Canadian Institutes of Health Research, Natural Sciences and Engineering Research Council of Canada as well as by Princess Margaret Hospital Foundation. We thank these funding agencies for supporting our research.

Conflicts of Interest: A.Z.-A. is a consultant with Point Surgical Inc.

References

1. Santos, C.R.; Schulze, A. Lipid metabolism in cancer. *FEBS J.* **2012**, *279*, 2610–2623. [CrossRef] [PubMed]
2. Zhang, J.; Sans, M.; Garza, K.Y.; Eberlin, L.S. Mass spectrometry technologies to advance care for cancer patients in clinical and intraoperative use. *Mass Spectrom. Rev.* **2020**, *40*, 692–720. [CrossRef] [PubMed]
3. Feider, C.L.; Krieger, A.C.; DeHoog, R.J.; Eberlin, L.S. Ambient Ionization Mass Spectrometry: Recent Developments and Applications. *Anal. Chem.* **2019**, *91*, 4266–4290. [CrossRef]
4. Takats, Z.; Strittmatter, N.; McKenzie, J. Ambient Mass Spectrometry in Cancer Research. *Methods Mol. Biol.* **2017**, *134*, 231–256. [CrossRef]
5. Ifa, D.R.; Eberlin, L.S. Ambient Ionization Mass Spectrometry for Cancer Diagnosis and Surgical Margin Evaluation. *Clin. Chem.* **2016**, *62*, 111–123. [CrossRef] [PubMed]
6. Clark, A.; Calligaris, D.; Regan, M.S.; Krummel, D.P.; Agar, J.N.; Kallay, L.; Macdonald, T.; Schniederjan, M.; Santagata, S.; Pomeroy, S.L.; et al. Rapid discrimination of pediatric brain tumors by mass spectrometry imaging. *J. Neuro-Oncol.* **2018**, *140*, 269–279. [CrossRef] [PubMed]
7. Eberlin, L.S. DESI-MS Imaging of Lipids and Metabolites from Biological Samples. In *Mass Spectrometry in Metabolomics*; Humana Press: New York, NY, USA, 2014; Volume 1198, pp. 299–311. [CrossRef]
8. Pirro, V.; Jarmusch, A.K.; Ferreira, C.R.; Cooks, R.G. Ambient Lipidomic Analysis of Brain Tissue Using Desorption Electrospray Ionization (DESI) Mass Spectrometry. *Neuromethods* **2017**, *125*, 187–210. [CrossRef]
9. Woolman, M.; Kuzan-Fischer, C.M.; Ferry, I.; Kiyota, T.; Luu, B.; Wu, M.; Munoz, D.G.; Das, S.; Aman, A.; Taylor, M.D.; et al. Picosecond Infrared Laser Desorption Mass Spectrometry Identifies Medulloblastoma Subgroups on Intrasurgical Timescales. *Cancer Res.* **2019**, *79*, 2426–2434. [CrossRef]
10. Santagata, S.; Eberlin, L.S.; Norton, I.; Calligaris, D.; Feldman, D.R.; Ide, J.L.; Liu, X.; Wiley, J.S.; Vestal, M.L.; Ramkissoon, S.H.; et al. Intraoperative mass spectrometry mapping of an onco-metabolite to guide brain tumor surgery. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 11121–11126. [CrossRef]
11. Calligaris, D.; Norton, I.; Feldman, D.R.; Ide, J.L.; Dunn, I.F.; Eberlin, L.S.; Cooks, R.G.; Jolesz, F.A.; Golby, A.J.; Santagata, S.; et al. Mass spectrometry imaging as a tool for surgical decision-making. *J. Mass Spectrom.* **2013**, *48*, 1178–1187. [CrossRef]

12. Eberlin, L.S.; Norton, I.; Orringer, D.; Dunn, I.F.; Liu, X.; Ide, J.L.; Jarmusch, A.K.; Ligon, K.L.; Jolesz, F.A.; Golby, A.J.; et al. Ambient mass spectrometry for the intraoperative molecular diagnosis of human brain tumors. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 1611–1616. [CrossRef]
13. Saudemont, P.; Quanico, J.; Robin, Y.-M.; Baud, A.; Balog, J.; Fatou, B.; Tierny, D.; Pascal, Q.; Minier, K.; Pottier, M.; et al. Real-Time Molecular Diagnosis of Tumors Using Water-Assisted Laser Desorption/Ionization Mass Spectrometry Technology. *Cancer Cell* **2018**, *34*, 840–851.e4. [CrossRef]
14. Balog, J.; Sasi-Szabó, L.; Kinross, J.; Lewis, M.R.; Muirhead, L.J.; Veselkov, K.; Mirnezami, R.; Dezso, B.; Damjanovich, L.; Darzi, A.; et al. Intraoperative Tissue Identification Using Rapid Evaporative Ionization Mass Spectrometry. *Sci. Transl. Med.* **2013**, *5*, 194ra93. [CrossRef]
15. Schäfer, K.-C.; Balog, J.; Szaniszló, T.; Szalay, D.; Mezey, G.; Dénes, J.; Bognár, L.; Oertel, M.; Takáts, Z. Real Time Analysis of Brain Tissue by Direct Combination of Ultrasonic Surgical Aspiration and Sonic Spray Mass Spectrometry. *Anal. Chem.* **2011**, *83*, 7729–7735. [CrossRef]
16. Schäfer, K.-C.; Szaniszló, T.; Günther, S.; Balog, J.; Dénes, J.; Keserű, M.; Dezso, B.; Tóth, M.; Spengler, B.; Takáts, Z. In Situ, Real-Time Identification of Biological Tissues by Ultraviolet and Infrared Laser Desorption Ionization Mass Spectrometry. *Anal. Chem.* **2011**, *83*, 1632–1640. [CrossRef]
17. Zhang, J.; Rector, J.; Lin, J.; Young, J.H.; Sans, M.; Katta, N.; Giese, N.; Yu, W.; Nagi, C.; Suliburk, J.; et al. Nondestructive tissue analysis for Ex Vivo and In Vivo cancer diagnosis using a handheld mass spectrometry system. *Sci. Transl. Med.* **2017**, *9*, eaan3968. [CrossRef]
18. Kiritani, S.; Yoshimura, K.; Arita, J.; Kokudo, T.; Hakoda, H.; Tanimoto, M.; Ishizawa, T.; Akamatsu, N.; Kaneko, J.; Takeda, S.; et al. A new rapid diagnostic system with ambient mass spectrometry and machine learning for colorectal liver metastasis. *BMC Cancer* **2021**, *21*, 262. [CrossRef] [PubMed]
19. Huang, Y.-C.; Chung, H.-H.; Dutkiewicz, E.P.; Chen, C.-L.; Hsieh, H.-Y.; Chen, B.-R.; Wang, M.-Y.; Hsu, C.-C. Predicting Breast Cancer by Paper Spray Ion Mobility Spectrometry Mass Spectrometry and Machine Learning. *Anal. Chem.* **2020**, *92*, 1653–1657. [CrossRef] [PubMed]
20. Zhou, Z.; Zare, R.N. Personal Information from Latent Fingerprints Using Desorption Electrospray Ionization Mass Spectrometry and Machine Learning. *Anal. Chem.* **2017**, *89*, 1369–1372. [CrossRef]
21. Tata, A.; Woolman, M.; Ventura, M.; Bernards, N.; Ganguly, M.; Gribble, A.; Shrestha, B.; Bluemke, E.; Ginsberg, H.J.; Vitkin, A.; et al. Rapid Detection of Necrosis in Breast Cancer with Desorption Electrospray Ionization Mass Spectrometry. *Sci. Rep.* **2016**, *6*, 35374. [CrossRef] [PubMed]
22. Koundouros, N.; Karali, E.; Tripp, A.; Valle, A.; Inglese, P.; Perry, N.J.; Magee, D.J.; Virmouni, S.A.; Elder, G.A.; Tyson, A.; et al. Metabolic Fingerprinting Links Oncogenic PIK3CA with Enhanced Arachidonic Acid-Derived Eicosanoids. *Cell* **2020**, *181*, 1596–1611.e27. [CrossRef]
23. Pirro, V.; Alfaro, C.M.; Jarmusch, A.K.; Hattab, E.M.; Cohen-Gadol, A.A.; Cooks, R.G. Intraoperative assessment of tumor margins during glioma resection by desorption electrospray ionization-mass spectrometry. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 6700–6705. [CrossRef] [PubMed]
24. Yannell, K.E.; Smith, K.; Alfaro, C.M.; Jarmusch, A.K.; Pirro, V.; Cooks, R.G. N-Acetylaspartate and 2-Hydroxyglutarate Assessed in Human Brain Tissue by Mass Spectrometry as Neuronal Markers of Oncogenesis. *Clin. Chem.* **2017**, *63*, 1766–1767. [CrossRef] [PubMed]
25. Jarmusch, A.K.; Pirro, V.; Baird, Z.; Hattab, E.; Cohen-Gadol, A.; Cooks, R.G. Lipid and metabolite profiles of human brain tumors by desorption electrospray ionization-MS. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 1486–1491. [CrossRef] [PubMed]
26. Reitman, Z.; Jin, G.; Karoly, E.D.; Spasojevic, I.; Yang, J.; Kinzler, K.W.; He, Y.; Bigner, D.D.; Vogelstein, B.; Yan, H. Profiling the effects of isocitrate dehydrogenase 1 and 2 mutations on the cellular metabolome. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 3270–3275. [CrossRef] [PubMed]
27. Eberlin, L.S.; Tibshirani, R.J.; Zhang, J.; Longacre, T.A.; Berry, G.J.; Bingham, D.B.; Norton, J.A.; Zare, R.N.; Poultsides, G.A. Molecular assessment of surgical-resection margins of gastric cancer by mass-spectrometric imaging. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 2436–2441. [CrossRef]
28. Chen, X.; Zhu, Y.; Jijiwa, M.; Nasu, M.; Ai, J.; Dai, S.; Jiang, B.; Zhang, J.; Huang, G.; Deng, Y. Identification of plasma lipid species as promising diagnostic markers for prostate cancer. *BMC Med. Inform. Decis. Mak.* **2020**, *20*, 223. [CrossRef]
29. Yan, F.; Zhao, H.; Zeng, Y. Lipidomics: A promising cancer biomarker. *Clin. Transl. Med.* **2018**, *7*, 21. [CrossRef]
30. Tata, A.; Woolman, M.; Bluemke, E.; Zarrine-Afsar, A. Chapter 5—Ambient laser-based mass spectrometry analysis methods: A survey of core technologies and reported applications. In *Ambient Ionization Mass Spectrometry in Life Sciences*; Zaitsu, K., Ed.; Elsevier: Amsterdam, The Netherlands, 2020; pp. 119–169.
31. Fatou, B.; Saudemont, P.; Leblanc, E.; Vinatier, D.; Mesdag, V.; Wisztorski, M.; Focsa, C.; Salzet, M.; Ziskind, M.; Fournier, I. In Vivo Real-Time Mass Spectrometry for Guided Surgery Application. *Sci. Rep.* **2016**, *6*, 25919. [CrossRef]
32. Woolman, M.; Ferry, I.; Kuzan-Fischer, C.M.; Wu, M.; Zou, J.; Kiyota, T.; Isik, S.; Dara, D.; Aman, A.; Das, S.; et al. Rapid determination of medulloblastoma subgroup affiliation with mass spectrometry using a handheld picosecond infrared laser desorption probe. *Chem. Sci.* **2017**, *8*, 6508–6519. [CrossRef]
33. Diamandis, E. The failure of protein cancer biomarkers to reach the clinic: Why, and what can be done to address the problem? *BMC Med.* **2012**, *10*, 87. [CrossRef] [PubMed]

34. Drucker, E.; Krapfenbauer, K. Pitfalls and limitations in translation from biomarker discovery to clinical utility in predictive and personalised medicine. *EPMA J.* **2013**, *4*, 7. [CrossRef] [PubMed]
35. Hanash, S.M. Why have protein biomarkers not reached the clinic? *Genome Med.* **2011**, *3*, 66. [CrossRef] [PubMed]
36. Raudys, S.; Jain, A.K. Small sample size effects in statistical pattern recognition: Recommendations for practitioners. *IEEE Trans. Pattern Anal. Mach. Intell.* **1991**, *13*, 252–264. [CrossRef]
37. Seddiki, K.; Saudemont, P.; Precioso, F.; Ogrinc, N.; Wisztorski, M.; Salzet, M.; Fournier, I.; Droit, A. Cumulative learning enables convolutional neural network representations for small mass spectrometry data classification. *Nat. Commun.* **2020**, *11*, 5595. [CrossRef] [PubMed]
38. Othman, R. Dietary Lipids and Cancer. *Libyan J. Med.* **2007**, *2*, 180–184. [CrossRef]
39. Peck, B.; Schulze, A. Lipid Metabolism at the Nexus of Diet and Tumor Microenvironment. *Trends Cancer* **2019**, *5*, 693–703. [CrossRef]
40. Xia, S.; Lin, R.; Jin, L.; Zhao, L.; Kang, H.-B.; Pan, Y.; Liu, S.; Qian, G.; Qian, Z.; Konstantakou, E.; et al. Prevention of Dietary-Fat-Fueled Ketogenesis Attenuates BRAF V600E Tumor Growth. *Cell Metab.* **2017**, *25*, 358–373. [CrossRef]
41. Dahdah, N.; Gonzalez-Franquesa, A.; Samino, S.; Gama-Perez, P.; Herrero, L.; Perales, J.; Yanes, O.; Malagón, M.; Garcia-Roves, P. Effects of Lifestyle Intervention in Tissue-Specific Lipidomic Profile of Formerly Obese Mice. *Int. J. Mol. Sci.* **2021**, *22*, 3694. [CrossRef]
42. Slenter, D.; Kutmon, M.; Hanspers, K.; Riutta, A.; Windsor, J.; Nunes, N.; Mélius, J.; Cirillo, E.; Coort, S.L.; Digles, D.; et al. WikiPathways: A multifaceted pathway database bridging metabolomics to other omics research. *Nucleic Acids Res.* **2018**, *46*, D661–D667. [CrossRef]
43. Park, J.K.; Coffey, N.J.; Limoges, A.; Le, A. The Heterogeneity of Lipid Metabolism in Cancer. *Adv. Exp. Med. Biol.* **2021**, *1311*, 39–56. [CrossRef]
44. Blaise, B.J.; Correia, G.D.S.; Tin, A.; Young, J.H.; Vergnaud, A.-C.; Lewis, M.; Pearce, J.; Elliott, P.; Nicholson, J.; Holmes, E.; et al. Power Analysis and Sample Size Determination in Metabolic Phenotyping. *Anal. Chem.* **2016**, *88*, 5179–5188. [CrossRef]
45. Blaise, B.J. Data-Driven Sample Size Determination for Metabolic Phenotyping Studies. *Anal. Chem.* **2013**, *85*, 8943–8950. [CrossRef]
46. Xia, J.; Broadhurst, D.; Wilson, M.; Wishart, D.S. Translational biomarker discovery in clinical metabolomics: An introductory tutorial. *Metabolomics* **2013**, *9*, 280–299. [CrossRef]
47. Guo, Y.; Graber, A.; McBurney, R.N.; Balasubramanian, R. Sample size and statistical power considerations in high-dimensionality data settings: A comparative study of classification algorithms. *BMC Bioinform.* **2010**, *11*, 447. [CrossRef]
48. Sun, T.; Wang, X.; Cong, P.; Xu, J.; Xue, C. Mass spectrometry-based lipidomics in food science and nutritional health: A comprehensive review. *Compr. Rev. Food Sci. Food Saf.* **2020**, *19*, 2530–2558. [CrossRef]
49. McGrath, T.F.; Haughey, S.A.; Patterson, J.; Fauhl-Hassek, C.; Donarski, J.; Alewijn, M.; van Ruth, S.; Elliott, C.T. What are the scientific challenges in moving from targeted to non-targeted methods for food fraud testing and how can they be addressed? Spectroscopy case study. *Trends Food Sci. Technol.* **2018**, *76*, 38–55. [CrossRef]
50. Esslinger, S.; Riedl, J.; Fauhl-Hassek, C. Potential and limitations of non-targeted fingerprinting for authentication of food in official control. *Food Res. Int.* **2014**, *60*, 189–204. [CrossRef]
51. Cavanna, D.; Righetti, L.; Elliott, C.; Suman, M. The scientific challenges in moving from targeted to non-targeted mass spectrometric methods for food fraud analysis: A proposed validation workflow to bring about a harmonized approach. *Trends Food Sci. Technol.* **2018**, *80*, 223–241. [CrossRef]
52. U.S. Pharmacopeia. *Appendix XVIII: Guidance on Developing and Validating Non-Targeted Methods for Adulteration Detection*; U.S. Pharmacopeial Convention: Rockville, MA, USA, 2018.
53. Marton, M.J.; Weiner, R. Practical Guidance for Implementing Predictive Biomarkers into Early Phase Clinical Studies. *BioMed Res. Int.* **2013**, *2013*, 891391. [CrossRef]
54. Ioannidis, J.P.A. Comparison of Effect Sizes Associated with Biomarkers Reported in Highly Cited Individual Articles and in Subsequent Meta-analyses. *JAMA* **2011**, *305*, 2200–2210. [CrossRef] [PubMed]
55. Ioannidis, J.P.A.; Bossuyt, P.M.M. Waste, Leaks, and Failures in the Biomarker Pipeline. *Clin. Chem.* **2017**, *63*, 963–972. [CrossRef] [PubMed]
56. Goossens, N.; Nakagawa, S.; Sun, X.; Hoshida, Y. Cancer biomarker discovery and validation. *Transl. Cancer Res.* **2015**, *4*, 256–269. [CrossRef]
57. Woolman, M.; Katz, L.; Tata, A.; Basu, S.S.; Zarrine-Afsar, A. Breaking Through the Barrier: Regulatory Considerations Relevant to Ambient Mass Spectrometry at the Bedside. *Clin. Lab. Med.* **2021**, *41*, 221–246. [CrossRef]
58. Kuo, T.-H.; Dutkiewicz, E.P.; Pei, J.; Hsu, C.-C. Ambient Ionization Mass Spectrometry Today and Tomorrow: Embracing Challenges and Opportunities. *Anal. Chem.* **2020**, *92*, 2353–2363. [CrossRef] [PubMed]
59. Draft Guidance for Industry, Food and Drug Administration Staff, and Clinical Laboratories. Available online: https://www.fda.gov/media/89841/download (accessed on 25 September 2021).
60. Kennedy, A.D.; Wittmann, B.M.; Evans, A.M.; Miller, L.A.; Toal, D.R.; Lonergan, S.; Elsea, S.H.; Pappan, K.L. Metabolomics in the clinic: A review of the shared and unique features of untargeted metabolomics for clinical research and clinical testing. *J. Mass Spectrom.* **2018**, *53*, 1143–1154. [CrossRef]

61. 510(k) Substantial Equivalence Determination Decision Summary, Number: K130831. Available online: https://www.accessdata.fda.gov/cdrh_docs/reviews/K130831.pdf (accessed on 25 September 2021).
62. 510(k) Substantial Equivalence Determination Decision Summary, Number: KK162950. Available online: https://www.accessdata.fda.gov/cdrh_docs/reviews/K162950.pdf (accessed on 25 September 2021).
63. USP Pharmacopeial Convention. *Appendix XVIII 2016: Guidance on Developing and Validating Nontargeted Methods for Adulteration Detection*; U.S. Pharmacopeial Convention: Rockville, MA, USA, 2016.
64. Bioanalytical Method Validation Guidance for Industry. Available online: https://www.fda.gov/files/drugs/published/Bioanalytical-Method-Validation-Guidance-for-Industry.pdf (accessed on 25 September 2021).
65. Franceschi, P.; Vrhovsek, U.; Mattivi, F.; Wehrens, R. *Metabolic Biomarker Identification with Few Samples*; IntechOpen: London, UK, 2012. [CrossRef]
66. Kreidler, S.M.; Muller, K.E.; Grunwald, G.K.; Ringham, B.M.; Coker-Dukowitz, Z.T.; Sakhadeo, U.R.; Baron, A.E.; Glueck, D.H. GLIMMPSE: Online Power Computation for Linear Models with and without a Baseline Covariate. *J. Stat. Softw.* **2013**, *54*, 1–26. [CrossRef]
67. Sumner, L.W.; Amberg, A.; Barrett, D.; Beale, M.H.; Beger, R.; Daykin, C.A.; Fan, T.W.-M.; Fiehn, O.; Goodacre, R.; Griffin, J.L.; et al. Proposed minimum reporting standards for chemical analysis. *Metabolomics* **2007**, *3*, 211–221. [CrossRef] [PubMed]
68. Katz, L.; Woolman, M.; Tata, A.; Zarrine-Afsar, A. Potential impact of tissue molecular heterogeneity on ambient mass spectrometry profiles: A note of caution in choosing the right disease model. *Anal. Bioanal. Chem.* **2021**, *413*, 2655–2664. [CrossRef]
69. Shevchenko, A.; Simons, K. Lipidomics: Coming to grips with lipid diversity. *Nat. Rev. Mol. Cell Biol.* **2010**, *11*, 593–598. [CrossRef]
70. Yang, K.; Han, X. Lipidomics: Techniques, Applications, and Outcomes Related to Biomedical Sciences. *Trends Biochem. Sci.* **2016**, *41*, 954–969. [CrossRef]
71. Milgraum, L.Z.; Witters, L.A.; Pasternack, G.R.; Kuhajda, F.P. Enzymes of the fatty acid synthesis pathway are highly expressed in In Situ breast carcinoma. *Clin. Cancer Res.* **1997**, *3*, 2115–2120.
72. Swinnen, J.V.; Vanderhoydonc, F.; Elgamal, A.A.; Eelen, M.; Vercaeren, I.; Joniau, S.; Van Poppel, H.; Baert, L.; Goossens, K.; Heyns, W.; et al. Selective activation of the fatty acid synthesis pathway in human prostate cancer. *Int. J. Cancer* **2000**, *88*, 176–179. [CrossRef]
73. Jung, Y.Y.; Kim, H.M.; Koo, J.S. Expression of Lipid Metabolism-Related Proteins in Metastatic Breast Cancer. *PLoS ONE* **2015**, *10*, e0137204. [CrossRef] [PubMed]
74. Avantaggiati, M.L. Cancer metabolism as a therapeutic target: Finding the right target(s) in the context of tumor heterogeneity, evolution, and metabolic plasticity. *Oncology* **2013**, *27*, 427–432.
75. Camarda, R.; Zhou, A.; Kohnz, R.A.; Balakrishnan, S.; Mahieu, C.; Anderton, B.; Eyob, H.; Kajimura, S.; Tward, A.; Krings, G.; et al. Inhibition of fatty acid oxidation as a therapy for MYC-overexpressing triple-negative breast cancer. *Nat. Med.* **2016**, *22*, 427–432. [CrossRef] [PubMed]
76. Dang, C.V.; Hamaker, M.; Sun, P.; Le, A.; Gao, P. Therapeutic targeting of cancer cell metabolism. *J. Mol. Med.* **2011**, *89*, 205–212. [CrossRef] [PubMed]
77. Ellsworth, R.E.; Blackburn, H.L.; Shriver, C.D.; Soon-Shiong, P.; Ellsworth, D.L. Molecular heterogeneity in breast cancer: State of the science and implications for patient care. *Semin. Cell Dev. Biol.* **2017**, *64*, 65–72. [CrossRef] [PubMed]
78. Cajal, S.R.Y.; Sesé, M.; Capdevila, C.; Aasen, T.; De Mattos-Arruda, L.; Diaz-Cano, S.J.; Hernández-Losa, J.; Castellvi, J. Clinical implications of intratumor heterogeneity: Challenges and opportunities. *J. Mol. Med.* **2020**, *98*, 161–177. [CrossRef]
79. Antonio, M.J.; Zhang, C.; Le, A. Different Tumor Microenvironments Lead to Different Metabolic Phenotypes. *Adv. Exp. Med. Biol.* **2021**, *1311*, 137–147. [CrossRef]
80. Nabi, K.; Le, A. The Intratumoral Heterogeneity of Cancer Metabolism. *Adv. Exp. Med. Biol.* **2021**, *1311*, 149–160. [CrossRef] [PubMed]
81. Salazar, J.; Le, A. The Heterogeneity of Liver Cancer Metabolism. *Adv. Exp. Med. Biol.* **2021**, *1311*, 127–136. [CrossRef] [PubMed]
82. Tan, J.; Le, A. The Heterogeneity of Breast Cancer Metabolism. *Adv. Exp. Med. Biol.* **2021**, *1311*, 89–101. [CrossRef]
83. Sharick, J.T.; Walsh, C.M.; Sprackling, C.M.; Pasch, C.A.; Pham, D.L.; Esbona, K.; Choudhary, A.; Varela, R.G.; Burkard, M.E.; McGregor, S.M.; et al. Metabolic Heterogeneity in Patient Tumor-Derived Organoids by Primary Site and Drug Treatment. *Front. Oncol.* **2020**, *10*, 553. [CrossRef] [PubMed]
84. Balaban, S.; Lee, L.S.; Varney, B.; Aishah, A.; Gao, Q.; Shearer, R.F.; Saunders, D.N.; Grewal, T.; Hoy, A.J. Heterogeneity of fatty acid metabolism in breast cancer cells underlies differential sensitivity to palmitate-induced apoptosis. *Mol. Oncol.* **2018**, *12*, 1623–1638. [CrossRef] [PubMed]
85. Gentric, G.; Mieulet, V.; Mechta-Grigoriou, F. Heterogeneity in Cancer Metabolism: New Concepts in an Old Field. *Antioxid. Redox Signal.* **2017**, *26*, 462–485. [CrossRef]
86. Hensley, C.T.; Faubert, B.; Yuan, Q.; Lev-Cohain, N.; Jin, E.; Kim, J.; Jiang, L.; Ko, B.; Skelton, R.; Loudat, L.; et al. Metabolic Heterogeneity in Human Lung Tumors. *Cell* **2016**, *164*, 681–694. [CrossRef] [PubMed]
87. Commander, R.; Wei, C.; Sharma, A.; Mouw, J.K.; Burton, L.J.; Summerbell, E.; Mahboubi, D.; Peterson, R.J.; Konen, J.; Zhou, W.; et al. Subpopulation targeting of pyruvate dehydrogenase and GLUT1 decouples metabolic heterogeneity during collective cancer cell invasion. *Nat. Commun.* **2020**, *11*, 2091–2105. [CrossRef]

88. Havas-Cavalletti, K.; Milchevskaya, V.; Radic, K.; Alladin, A.; Kafkia, E.; Garcia, M.; Stolte, J.; Klaus, B.; Rotmensz, N.; Gibson, T.J.; et al. Metabolic shifts in residual breast cancer drive tumor recurrence. *J. Clin. Investig.* **2017**, *127*, 2091–2105. [CrossRef]
89. Kim, J.; DeBerardinis, R.J. Mechanisms and Implications of Metabolic Heterogeneity in Cancer. *Cell Metab.* **2019**, *30*, 434–446. [CrossRef]
90. Benfeitas, R.; Uhlen, M.; Nielsen, J.; Mardinoglu, A. New Challenges to Study Heterogeneity in Cancer Redox Metabolism. *Front. Cell Dev. Biol.* **2017**, *5*, 65. [CrossRef] [PubMed]
91. Angerer, T.B.; Magnusson, Y.; Landberg, G.; Fletcher, J.S. Lipid Heterogeneity Resulting from Fatty Acid Processing in the Human Breast Cancer Microenvironment Identified by GCIB-ToF-SIMS Imaging. *Anal. Chem.* **2016**, *88*, 11946–11954. [CrossRef] [PubMed]
92. Lv, J.; Gao, D.; Zhang, Y.; Wu, D.; Shen, L.; Wang, X. Heterogeneity of lipidomic profiles among lung cancer subtypes of patients. *J. Cell. Mol. Med.* **2018**, *22*, 5155–5159. [CrossRef]
93. Li, Z.; Cheng, S.; Lin, Q.; Cao, W.; Yang, J.; Zhang, M.; Shen, A.; Zhang, W.; Xia, Y.; Ma, X.; et al. Single-cell lipidomics with high structural specificity by mass spectrometry. *Nat. Commun.* **2021**, *12*, 1237–1250. [CrossRef]
94. Blomme, A.; Van Simaeys, G.; Doumont, G.; Costanza, B.; Bellier, J.; Otaka, Y.; Sherer, F.; Lovinfosse, P.; Boutry, S.; Palacios, A.P.; et al. Murine stroma adopts a human-like metabolic phenotype in the PDX model of colorectal cancer and liver metastases. *Oncogene* **2018**, *37*, 1237–1250. [CrossRef]
95. Agar, N.Y.; Golby, A.J.; Ligon, K.L.; Norton, I.; Mohan, V.; Wiseman, J.M.; Tannenbaum, A.; Jolesz, F.A. Development of Stereotactic Mass Spectrometry for Brain Tumor Surgery. *Neurosurgery* **2011**, *68*, 280–290. [CrossRef]
96. Brown, H.M.; Alfaro, C.M.; Pirro, V.; Dey, M.; Hattab, E.M.; Cohen-Gadol, A.A.; Cooks, R.G. Intraoperative Mass Spectrometry Platform for IDH Mutation Status Prediction, Glioma Diagnosis, and Estimation of Tumor Cell Infiltration. *J. Appl. Lab. Med.* **2021**, *6*, 902–916. [CrossRef]
97. Wolff, A.; Hammond, M.E.H.; Hicks, D.G.; Dowsett, M.; McShane, L.M.; Allison, K.H.; Allred, D.C.; Bartlett, J.M.; Bilous, M.; Fitzgibbons, P.; et al. Recommendations for Human Epidermal Growth Factor Receptor 2 Testing in Breast Cancer: American Society of Clinical Oncology/College of American Pathologists Clinical Practice Guideline Update. *Arch. Pathol. Lab. Med.* **2014**, *138*, 241–256. [CrossRef]
98. Rauser, S.; Marquardt, C.; Balluff, B.; Deininger, S.-O.; Albers, C.; Belau, E.; Hartmer, R.; Suckau, D.; Specht, K.; Ebert, M.P.; et al. Classification of HER2 Receptor Status in Breast Cancer Tissues by MALDI Imaging Mass Spectrometry. *J. Proteome Res.* **2010**, *9*, 1854–1863. [CrossRef] [PubMed]
99. Porcari, A.; Zhang, J.; Garza, K.Y.; Peres, R.M.R.; Lin, J.; Young, J.H.; Tibshirani, R.; Nagi, C.; Paiva, G.R.; Carter, S.A.; et al. Multicenter Study Using Desorption-Electrospray-Ionization-Mass-Spectrometry Imaging for Breast-Cancer Diagnosis. *Anal. Chem.* **2018**, *90*, 11324–11332. [CrossRef] [PubMed]
100. Buckley, N.; Forde, C.; McArt, D.G.; Boyle, D.P.; Mullan, P.B.; James, J.A.; Maxwell, P.; McQuaid, S.; Salto-Tellez, M. Quantification of HER2 heterogeneity in breast cancer–implications for identification of sub-dominant clones for personalised treatment. *Sci. Rep.* **2016**, *6*, 23383. [CrossRef]
101. Sans, M.; Gharpure, K.; Tibshirani, R.; Zhang, J.; Liang, L.; Liu, J.; Young, J.H.; Dood, R.L.; Sood, A.K.; Eberlin, L.S. Metabolic Markers and Statistical Prediction of Serous Ovarian Cancer Aggressiveness by Ambient Ionization Mass Spectrometry Imaging. *Cancer Res.* **2017**, *77*, 2903–2913. [CrossRef]
102. Thompson, D.O.; Karmali, R.A.; Cohen, L.A.; Choi, K.; Rose, D.P. Dietary Fat and Mammary Cancer. II. Modulation of Serum and Tumor Lipid Composition and Tumor Prostaglandins by Different Dietary Fats: Association with Tumor Incidence Patterns. *J. Natl. Cancer Inst.* **1986**, *77*, 43–51. [CrossRef]
103. O'Keefe, S.J.D.; Li, J.; Lahti, L.; Ou, J.; Carbonero, F.; Mohammed, K.; Posma, J.M.; Kinross, J.; Wahl, E.; Ruder, E.; et al. Fat, fibre and cancer risk in African Americans and rural Africans. *Nat. Commun.* **2015**, *6*, 6342. [CrossRef]
104. Ellis, J.K.; Athersuch, T.J.; Thomas, L.D.K.; Teichert, F.; Pérez-Trujillo, M.; Svendsen, C.; Spurgeon, D.J.; Singh, R.; Järup, L.; Bundy, J.G.; et al. Metabolic profiling detects early effects of environmental and lifestyle exposure to cadmium in a human population. *BMC Med.* **2012**, *10*, 61. [CrossRef] [PubMed]
105. Finicle, B.T.; Jayashankar, V.; Edinger, A.L. Nutrient scavenging in cancer. *Nat. Rev. Cancer* **2018**, *18*, 619–633. [CrossRef] [PubMed]
106. Bruce, K.D.; Zsombok, A.; Eckel, R.H. Lipid Processing in the Brain: A Key Regulator of Systemic Metabolism. *Front. Endocrinol.* **2017**, *8*, 60. [CrossRef]
107. Woolman, M.; Katz, L.; Gopinath, G.; Kiyota, T.; Kuzan-Fischer, C.M.; Ferry, I.; Zaidi, M.; Peters, K.; Aman, A.; McKee, T.; et al. Mass Spectrometry Imaging Reveals a Gradient of Cancer-like Metabolic States in the Vicinity of Cancer Not Seen in Morphometric Margins from Microscopy. *Anal. Chem.* **2021**, *93*, 4408–4416. [CrossRef]
108. Sasaki, A.; Nagatake, T.; Egami, R.; Gu, G.; Takigawa, I.; Ikeda, W.; Nakatani, T.; Kunisawa, J.; Fujita, Y. Obesity Suppresses Cell-Competition-Mediated Apical Elimination of RasV12-Transformed Cells from Epithelial Tissues. *Cell Rep.* **2018**, *23*, 974–982. [CrossRef] [PubMed]
109. Evans, K.; Burdge, G.C.; Wootton, S.A.; Clark, M.L.; Frayn, K.N. Regulation of dietary fatty acid entrapment in subcutaneous adipose tissue and skeletal muscle. *Diabetes* **2002**, *51*, 2684–2690. [CrossRef] [PubMed]
110. Lengyel, E.; Makowski, L.; DiGiovanni, J.; Kolonin, M.G. Cancer as a Matter of Fat: The Crosstalk between Adipose Tissue and Tumors. *Trends Cancer* **2018**, *4*, 374–384. [CrossRef]
111. Zhao, C.; Wu, M.; Zeng, N.; Xiong, M.; Hu, W.; Lv, W.; Yi, Y.; Zhang, Q.; Wu, Y. Cancer-associated adipocytes: Emerging supporters in breast cancer. *J. Exp. Clin. Cancer Res.* **2020**, *39*, 156. [CrossRef] [PubMed]

112. Yang, D.; Li, Y.; Xing, L.; Tan, Y.; Sun, J.; Zeng, B.; Xiang, T.; Tan, J.; Ren, G.; Wang, Y. Utilization of adipocyte-derived lipids and enhanced intracellular trafficking of fatty acids contribute to breast cancer progression. *Cell Commun. Signal.* **2018**, *16*, 32. [CrossRef] [PubMed]
113. Volden, P.A.; Skor, M.N.; Johnson, M.B.; Singh, P.; Patel, F.N.; McClintock, M.K.; Brady, M.J.; Conzen, S.D. Mammary Adipose Tissue-Derived Lysophospholipids Promote Estrogen Receptor–Negative Mammary Epithelial Cell Proliferation. *Cancer Prev. Res.* **2016**, *9*, 367–378. [CrossRef] [PubMed]
114. Attané, C.; Milhas, D.; Hoy, A.; Muller, C. Metabolic Remodeling Induced by Adipocytes: A New Achilles' Heel in Invasive Breast Cancer? *Curr. Med. Chem.* **2020**, *27*, 3984–4001. [CrossRef]
115. Munteanu, R.; Onaciu, A.; Moldovan, C.; Zimta, A.-A.; Gulei, D.; Paradiso, A.V.; Lazar, V.; Berindan-Neagoe, I. Adipocyte-Based Cell Therapy in Oncology: The Role of Cancer-Associated Adipocytes and Their Reinterpretation as Delivery Platforms. *Pharmaceutics* **2020**, *12*, 402. [CrossRef] [PubMed]
116. Zhong, W.; Zhao, Y.; Tang, Y.; Wei, X.; Shi, X.; Sun, W.; Sun, X.; Yin, X.; Sun, X.; Kim, S.; et al. Chronic Alcohol Exposure Stimulates Adipose Tissue Lipolysis in Mice: Role of Reverse Triglyceride Transport in the Pathogenesis of Alcoholic Steatosis. *Am. J. Pathol.* **2012**, *180*, 998–1007. [CrossRef]
117. Pravdova, E.; Fickova, M. Alcohol intake modulates hormonal activity of adipose tissue. *Endocr. Regul.* **2006**, *40*, 91–104.
118. Wei, X.; Shi, X.; Zhong, W.; Zhao, Y.; Tang, Y.; Sun, W.; Yin, X.; Bogdanov, B.; Kim, S.; McClain, C.; et al. Chronic Alcohol Exposure Disturbs Lipid Homeostasis at the Adipose Tissue-Liver Axis in Mice: Analysis of Triacylglycerols Using High-Resolution Mass Spectrometry in Combination with In Vivo Metabolite Deuterium Labeling. *PLoS ONE* **2013**, *8*, e55382. [CrossRef]
119. Sales, S.; Graessler, J.; Ciucci, S.; Al-Atrib, R.; Vihervaara, T.; Schuhmann, K.; Kauhanen, D.; Sysi-Aho, M.; Bornstein, S.R.; Bickle, M.; et al. Gender, Contraceptives and Individual Metabolic Predisposition Shape a Healthy Plasma Lipidome. *Sci. Rep.* **2016**, *6*, 27710. [CrossRef]
120. Slade, E.; Irvin, M.R.; Xie, K.; Arnett, D.K.; Claas, S.A.; Kind, T.; Fardo, D.W.; Graf, G.A. Age and sex are associated with the plasma lipidome: Findings from the GOLDN study. *Lipids Health Dis.* **2021**, *20*, 30. [CrossRef] [PubMed]
121. Audano, M.; Maldini, M.; De Fabiani, E.; Mitro, N.; Caruso, D. Gender-related metabolomics and lipidomics: From experimental animal models to clinical evidence. *J. Proteom.* **2018**, *178*, 82–91. [CrossRef] [PubMed]
122. Jain, A.; Chandrasekaran, B. 39 Dimensionality and sample size considerations in pattern recognition practice. In *Handbook of Statistics*; Elsevier: Amsterdam, The Netherlands, 1982; Volume 2, pp. 835–855. [CrossRef]
123. Kalayeh, H.M.; Landgrebe, D.A. Predicting the Required Number of Training Samples. *IEEE Trans. Pattern Anal. Mach. Intell.* **1983**, *PAMI-5*, 664–667. [CrossRef] [PubMed]
124. Mukherjee, S.; Tamayo, P.; Rogers, S.; Rifkin, R.; Engle, A.; Campbell, C.; Golub, T.R.; Mesirov, J.P. Estimating Dataset Size Requirements for Classifying DNA Microarray Data. *J. Comput. Biol.* **2003**, *10*, 119–142. [CrossRef] [PubMed]
125. Dobbin, K.K.; Zhao, Y.; Simon, R.M. How Large a Training Set is Needed to Develop a Classifier for Microarray Data? *Clin. Cancer Res.* **2008**, *14*, 108–114. [CrossRef] [PubMed]
126. Figueroa, R.L.; Zeng-Treitler, Q.; Kandula, S.; Ngo, L.H. Predicting sample size required for classification performance. *BMC Med. Inform. Decis. Mak.* **2012**, *12*, 8. [CrossRef] [PubMed]
127. Beleites, C.; Neugebauer, U.; Bocklitz, T.; Krafft, C.; Popp, J. Sample size planning for classification models. *Anal. Chim. Acta* **2013**, *760*, 25–33. [CrossRef]
128. Lee, P.Y.; Yeoh, Y.; Omar, N.; Pung, Y.-F.; Lim, L.C.; Low, T.Y. Molecular tissue profiling by MALDI imaging: Recent progress and applications in cancer research. *Crit. Rev. Clin. Lab. Sci.* **2021**, 1–17. [CrossRef]
129. Mascini, N.E.; Teunissen, J.; Noorlag, R.; Willems, S.M.; Heeren, R.M. Tumor classification with MALDI-MSI data of tissue microarrays: A case study. *Methods* **2018**, *151*, 21–27. [CrossRef]
130. Klein, O.; Kanter, F.; Kulbe, H.; Jank, P.; Denkert, C.; Nebrich, G.; Schmitt, W.; Wu, Z.; Kunze, C.A.; Sehouli, J.; et al. MALDI-Imaging for Classification of Epithelial Ovarian Cancer Histotypes from a Tissue Microarray Using Machine Learning Methods. *Proteom. Clin. Appl.* **2019**, *13*, e1700181. [CrossRef]
131. Kassuhn, W.; Klein, O.; Darb-Esfahani, S.; Lammert, H.; Handzik, S.; Taube, E.; Schmitt, W.; Keunecke, C.; Horst, D.; Dreher, F.; et al. Classification of Molecular Subtypes of High-Grade Serous Ovarian Cancer by MALDI-Imaging. *Cancers* **2021**, *13*, 1512. [CrossRef] [PubMed]
132. Cairns, D.A.; Barrett, J.; Billingham, L.; Stanley, A.J.; Xinarianos, G.; Field, J.; Johnson, P.J.; Selby, P.J.; Banks, R. Sample size determination in clinical proteomic profiling experiments using mass spectrometry for class comparison. *Proteomics* **2009**, *9*, 74–86. [CrossRef] [PubMed]
133. Demidenko, E. Sample size determination for logistic regression revisited. *Stat. Med.* **2007**, *26*, 3385–3397. [CrossRef] [PubMed]
134. Smyth, G.K.; Michaud, J.; Scott, H. Use of within-array replicate spots for assessing differential expression in microarray experiments. *Bioinformatics* **2005**, *21*, 2067–2075. [CrossRef] [PubMed]
135. Nyangoma, S.O.; Collins, S.I.; Altman, D.G.; Johnson, P.; Billingham, L. Sample Size Calculations for Designing Clinical Proteomic Profiling Studies Using Mass Spectrometry. *Stat. Appl. Genet. Mol. Biol.* **2012**, *11*. [CrossRef]
136. Alfaro, C.M.; Jarmusch, A.K.; Pirro, V.; Kerian, K.S.; Masterson, T.A.; Cheng, L.; Cooks, R.G. Ambient ionization mass spectrometric analysis of human surgical specimens to distinguish renal cell carcinoma from healthy renal tissue. *Anal. Bioanal. Chem.* **2016**, *408*, 5407–5414. [CrossRef]

137. Margulis, K.; Chiou, A.S.; Aasi, S.Z.; Tibshirani, R.J.; Tang, J.Y.; Zare, R.N. Distinguishing malignant from benign microscopic skin lesions using desorption electrospray ionization mass spectrometry imaging. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 6347–6352. [CrossRef] [PubMed]
138. Giordano, S.; Takeda, S.; Donadon, M.; Saiki, H.; Brunelli, L.; Pastorelli, R.; Cimino, M.; Soldani, C.; Franceschini, B.; Di Tommaso, L.; et al. Rapid automated diagnosis of primary hepatic tumour by mass spectrometry and artificial intelligence. *Liver Int.* **2020**, *40*, 3117–3124. [CrossRef]
139. Paraskevaidi, M.; Cameron, S.J.; Whelan, E.; Bowden, S.; Tzafetas, M.; Mitra, A.; Semertzidou, A.; Athanasiou, A.; Bennett, P.R.; MacIntyre, D.A.; et al. Laser-assisted rapid evaporative ionisation mass spectrometry (LA-REIMS) as a metabolomics platform in cervical cancer screening. *EBioMedicine* **2020**, *60*, 103017. [CrossRef]
140. Van Kersbergen, J.; Ghazvinian Zanjani, F.; Zinger, S.; Van der Sommen, F.; Balluff, B.; Vos, N.; Ellis, S.; Heeran, R.; Lucas, M.; Marquering, H.; et al. Cancer detection in mass spectrometry imaging data by dilated convolutional neural networks. In Proceedings of the 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019), Venice, Italy, 8–11 April 2019; p. 16.
141. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. ImageNet classification with deep convolutional neural networks. *Commun. ACM* **2012**, *25*, 84–90. [CrossRef]
142. Forman, G.; Cohen, I. Learning from little: Comparison of classifiers given little training. In *European Conference on Principles of Data Mining and Knowledge Discovery*; Springer: Berlin/Heidelberg, Germany, 2004; Volume 3202.
143. Shaikhina, T.; Khovanova, N.A. Handling limited datasets with neural networks in medical applications: A small-data approach. *Artif. Intell. Med.* **2017**, *75*, 51–63. [CrossRef]
144. Bodai, Z.; Cameron, S.; Bolt, F.; Simon, D.; Schaffer, R.; Karancsi, T.; Balog, J.; Rickards, T.; Burke, A.; Hardiman, K.; et al. Effect of Electrode Geometry on the Classification Performance of Rapid Evaporative Ionization Mass Spectrometric (REIMS) Bacterial Identification. *J. Am. Soc. Mass Spectrom.* **2018**, *29*, 26–33. [CrossRef]
145. Johnson, C.; Gonzalez, F.J. Challenges and opportunities of metabolomics. *J. Cell. Physiol.* **2012**, *227*, 2975–2981. [CrossRef]
146. Johnson, C.; Ivanisevic, J.; Siuzdak, G. Metabolomics: Beyond biomarkers and towards mechanisms. *Nat. Rev. Mol. Cell Biol.* **2016**, *17*, 451–459. [CrossRef]
147. Simon, I.; Zhuo, S.; Corral, L.; Diamandis, E.; Sarno, M.J.; Wolfert, R.L.; Kim, N.W. B7-H4 Is a Novel Membrane-Bound Protein and a Candidate Serum and Tissue Biomarker for Ovarian Cancer. *Cancer Res.* **2006**, *66*, 1570–1575. [CrossRef]
148. Cramer, D.W.; Bast, R.; Berg, C.D.; Diamandis, E.; Godwin, A.K.; Hartge, P.; Lokshin, A.E.; Lu, K.H.; McIntosh, M.W.; Mor, G.; et al. Ovarian Cancer Biomarker Performance in Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial Specimens. *Cancer Prev. Res.* **2011**, *4*, 365–374. [CrossRef]
149. Xu, Y.; Shen, Z.; Wiper, D.W.; Wu, M.; Morton, R.E.; Elson, P.; Kennedy, A.W.; Belinson, J.; Markman, M.; Casey, G. Lysophosphatidic Acid as a Potential Biomarker for Ovarian and Other Gynecologic Cancers. *JAMA* **1998**, *280*, 719–723. [CrossRef] [PubMed]
150. Baker, D.L. Plasma Lysophosphatidic Acid Concentration and Ovarian Cancer. *JAMA* **2002**, *287*, 3081–3082. [CrossRef]
151. Stephan, C.; Wilkosz, J.; Różański, W.; Ecke, T.H.; Lein, M.; Bryś, M.; Krześlak, A.; Chwatko, G.; Jung, K. Urinary thiosulfate as failed prostate cancer biomarker—An exemplary multicenter re-evaluation study. *Clin. Chem. Lab. Med.* **2015**, *53*, 477–483. [CrossRef] [PubMed]
152. Massaro, A.; Negro, A.; Bragolusi, M.; Miano, B.; Tata, A.; Suman, M.; Piro, R. Oregano authentication by mid-level data fusion of chemical fingerprint signatures acquired by ambient mass spectrometry. *Food Control* **2021**, *126*, 108058. [CrossRef]
153. Damiani, T.; Cavanna, D.; Serani, A.; Dall'Asta, C.; Suman, M. GC-IMS and FGC-Enose fingerprint as screening tools for revealing extra virgin olive oil blending with soft-refined olive oils: A feasibility study. *Microchem. J.* **2020**, *159*, 105374. [CrossRef]
154. Kunz, M.; Ottaway, J.; Kalivas, J.H.; Georgiou, C.; Mousdis, G. Updating a Synchronous Fluorescence Spectroscopic Virgin Olive Oil Adulteration Calibration to a New Geographical Region. *J. Agric. Food Chem.* **2011**, *59*, 1051–1057. [CrossRef]
155. Riedl, J.; Esslinger, S.; Fauhl-Hassek, C. Review of validation and reporting of non-targeted fingerprinting approaches for food authentication. *Anal. Chim. Acta* **2015**, *885*, 17–32. [CrossRef] [PubMed]
156. Schymanski, E.; Jeon, J.; Gulde, R.; Fenner, K.; Ruff, M.; Singer, H.; Hollender, J. Identifying Small Molecules via High Resolution Mass Spectrometry: Communicating Confidence. *Environ. Sci. Technol.* **2014**, *48*, 2097–2098. [CrossRef] [PubMed]
157. Nishi, A.; Ohbuchi, K.; Kaifuchi, N.; Shimobori, C.; Kushida, H.; Yamamoto, M.; Kita, Y.; Tokuoka, S.M.; Yachie, A.; Matsuoka, Y.; et al. Publisher Correction: LimeMap: A comprehensive map of lipid mediator metabolic pathways. *NPJ Syst. Biol. Appl.* **2021**, *7*, 16. [CrossRef] [PubMed]
158. Nishi, A.; Ohbuchi, K.; Kaifuchi, N.; Shimobori, C.; Kushida, H.; Yamamoto, M.; Kita, Y.; Tokuoka, S.M.; Yachie, A.; Matsuoka, Y.; et al. LimeMap: A comprehensive map of lipid mediator metabolic pathways. *NPJ Syst. Biol. Appl.* **2021**, *7*, 6. [CrossRef]
159. Johnson, C.; Ivanisevic, J.; Benton, H.P.; Siuzdak, G. Bioinformatics: The Next Frontier of Metabolomics. *Anal. Chem.* **2015**, *87*, 147–156. [CrossRef] [PubMed]
160. Thiele, I.; Swainston, N.; Fleming, R.M.; Hoppe, A.; Sahoo, S.; Aurich, M.K.; Haraldsdottir, H.; Mo, M.L.; Rolfsson, O.; Stobbe, M.D.; et al. A community-driven global reconstruction of human metabolism. *Nat. Biotechnol.* **2013**, *31*, 419–425. [CrossRef]
161. Zamboni, N.; Saghatelian, A.; Patti, G.J. Defining the Metabolome: Size, Flux, and Regulation. *Mol. Cell* **2015**, *58*, 699–706. [CrossRef]

162. Li, S.; Park, Y.; Duraisingham, S.; Strobel, F.H.; Khan, N.; Soltow, Q.A.; Jones, D.P.; Pulendran, B. Predicting Network Activity from High Throughput Metabolomics. *PLoS Comput. Biol.* **2013**, *9*, e1003123. [CrossRef]
163. Cottret, L.; Wildridge, D.; Vinson, F.; Barrett, M.; Charles, H.; Sagot, M.-F.; Jourdan, F. MetExplore: A web server to link metabolomic experiments and genome-scale metabolic networks. *Nucleic Acids Res.* **2010**, *38*, W132–W137. [CrossRef]
164. Scalbert, A.; Brennan, L.; Manach, C.; Andres-Lacueva, C.; Dragsted, L.O.; Draper, J.; Rappaport, S.M.; van der Hooft, J.J.; Wishart, D.S. The food metabolome: A window over dietary exposure. *Am. J. Clin. Nutr.* **2014**, *99*, 1286–1308. [CrossRef]
165. Wishart, D.S.; Jewison, T.; Guo, A.C.; Wilson, M.; Knox, C.; Liu, Y.; Djoumbou, Y.; Mandal, R.; Aziat, F.; Dong, E.; et al. HMDB 3.0—The Human Metabolome Database in 2013. *Nucleic Acids Res.* **2013**, *41*, D801–D807. [CrossRef]
166. Tautenhahn, R.; Patti, G.J.; Rinehart, D.; Siuzdak, G. XCMS Online: A Web-Based Platform to Process Untargeted Metabolomic Data. *Anal. Chem.* **2012**, *84*, 5035–5039. [CrossRef]
167. Xia, J.; Psychogios, N.; Young, N.; Wishart, D.S. MetaboAnalyst: A web server for metabolomic data analysis and interpretation. *Nucleic Acids Res.* **2009**, *37*, W652–W660. [CrossRef]
168. Margulis, K.; Zhou, Z.; Fang, Q.; Sievers, R.E.; Lee, R.J.; Zare, R.N. Combining Desorption Electrospray Ionization Mass Spectrometry Imaging and Machine Learning for Molecular Recognition of Myocardial Infarction. *Anal. Chem.* **2018**, *90*, 12198–12206. [CrossRef]
169. Schaffer, S.W.; Jong, C.J.; Kc, R.; Azuma, J. Physiological roles of taurine in heart and muscle. *J. Biomed. Sci.* **2010**, *17* (Suppl. 1), S2–S8. [CrossRef] [PubMed]
170. Schaffer, S.W.; Jong, C.J.; Ito, T.; Azuma, J. Effect of taurine on ischemia-reperfusion injury. *Amino Acids* **2014**, *46*, 21–30. [CrossRef] [PubMed]
171. Ardisson, L.P.; Rafacho, B.P.; Santos, P.P.; Assalin, H.; Gonçalves, A.F.; Azevedo, P.S.; Minicucci, M.F.; Polegato, B.F.; Okoshi, K.; Marchini, J.S.; et al. Taurine attenuates cardiac remodeling after myocardial infarction. *Int. J. Cardiol.* **2013**, *168*, 4925–4926. [CrossRef] [PubMed]
172. Banerjee, S.; Zare, R.N.; Tibshirani, R.J.; Kunder, C.; Nolley, R.; Fan, R.; Brooks, J.D.; Sonn, G.A. Diagnosis of prostate cancer by desorption electrospray ionization mass spectrometric imaging of small metabolites and lipids. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 3334–3339. [CrossRef] [PubMed]
173. Liebisch, G.; Fahy, E.; Aoki, J.; Dennis, E.A.; Durand, T.; Ejsing, C.S.; Fedorova, M.; Feussner, I.; Griffiths, W.J.; Köfeler, H.; et al. Update on LIPID MAPS classification, nomenclature, and shorthand notation for MS-derived lipid structures. *J. Lipid Res.* **2020**, *61*, 1539–1555. [CrossRef] [PubMed]
174. Liebisch, G.; Vizcaino, J.A.; Köfeler, H.; Trötzmüller, M.; Griffiths, W.; Schmitz, G.; Spener, F.; Wakelam, M. Shorthand notation for lipid structures derived from mass spectrometry. *J. Lipid Res.* **2013**, *54*, 1523–1530. [CrossRef] [PubMed]
175. Lipidomics Standards Initiative Consortium. Lipidomics needs more standardization. *Nat. Metab.* **2019**, *1*, 745–747. [CrossRef] [PubMed]