



HAL
open science

Vitacrytallography: Structural Biomarkers of Breast Cancer Obtained by X-ray Scattering

Sergey Denisov, Benjamin Blinchevsky, Jonathan Friedman, Barbara Gerbelli, Ash Ajeer, Lois Adams, Charlene Greenwood, Keith Rogers, Lev Mourokh, Pavel Lazarev

► **To cite this version:**

Sergey Denisov, Benjamin Blinchevsky, Jonathan Friedman, Barbara Gerbelli, Ash Ajeer, et al.. Vitacrytallography: Structural Biomarkers of Breast Cancer Obtained by X-ray Scattering. *Cancers*, 2024, 16 (14), pp.2499. <10.3390/cancers16142499>. <hal-05293597>

HAL Id: hal-05293597

<https://hal.science/hal-05293597v1>

Submitted on 2 Oct 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY 4.0 - Attribution - International License

Article

Vitacrystallography: Structural Biomarkers of Breast Cancer Obtained by X-ray Scattering

Sergey Denisov^{1,2}, Benjamin Blinchevsky^{1,3}, Jonathan Friedman^{3,4}, Barbara Gerbelli⁵, Ash Ajeer⁵ , Lois Adams⁵, Charlene Greenwood^{3,5}, Keith Rogers^{3,6} , Lev Mourokh^{3,4,*} and Pavel Lazarev^{1,3}

¹ Matur UK Ltd., 5 New Street Square, London EC4A 3TW, UK; sdenisov@matur.co.uk (S.D.); bblinchevsky@matur.co.uk (B.B.); plazarev@matur.co.uk (P.L.)

² Institut de Chimie Physique, UMR8000, CNRS, Université Paris-Saclay, Bât. 349, 91405 Orsay, France

³ EosDx, Inc., 1455 Adams Drive, Menlo Park, CA 94025, USA; jonbenfri@gmail.com (J.F.); c.e.greenwood@keele.ac.uk (C.G.); k.d.rogers@cranfield.ac.uk (K.R.)

⁴ Physics Department, Queens College, City University of New York, 65-30 Kissena Blvd, Flushing, NY 11367, USA

⁵ School of Chemical and Physical Sciences, Keele University, Keele ST5 5BG, UK; barbara.gerbelli@diamond.ac.uk (B.G.); a.ajeer@keele.ac.uk (A.A.); l.e.m.adams@keele.ac.uk (L.A.)

⁶ Shrivenham Campus, Cranfield University, Swindon SN6 8LA, UK

* Correspondence: lev.murokh@qc.cuny.edu

Simple Summary: Breast cancer ranks as the most prevalent cancer among women. Current screening includes regular mammography and subsequent biopsy if the mammography results are abnormal. These procedures are costly and uncomfortable. We propose an alternative non-invasive method based on X-ray scattering. Using a machine learning approach, we have examined almost 3000 measurements of cancerous and non-cancerous samples belonging to 110 patients and shown excellent results on cancer/non-cancer separation. This can lead to patient-friendly, fast, and economical solutions for breast cancer screening to complement mammography and reduce biopsy. It should be emphasized that this approach can be readily extended to other types of cancer and even other diseases.



Citation: Denisov, S.; Blinchevsky, B.; Friedman, J.; Gerbelli, B.; Ajeer, A.; Adams, L.; Greenwood, C.; Rogers, K.; Mourokh, L.; Lazarev, P.

Vitacrystallography: Structural Biomarkers of Breast Cancer Obtained by X-ray Scattering. *Cancers* **2024**, *16*, 2499. <https://doi.org/10.3390/cancers16142499>

Academic Editor: Hans Knecht

Received: 27 May 2024

Revised: 26 June 2024

Accepted: 8 July 2024

Published: 9 July 2024

Abstract: With breast cancer being one of the most widespread causes of death for women, there is an unmet need for its early detection. For this purpose, we propose a non-invasive approach based on X-ray scattering. We measured samples from 107 unique patients provided by the Breast Cancer Now Tissue Biobank, with the total dataset containing 2958 entries. Two different sample-to-detector distances, 2 and 16 cm, were used to access various structural biomarkers at distinct ranges of momentum transfer values. The biomarkers related to lipid metabolism are consistent with those of previous studies. Machine learning analysis based on the Random Forest Classifier demonstrates excellent performance metrics for cancer/non-cancer binary decisions. The best sensitivity and specificity values are 80% and 92%, respectively, for the sample-to-detector distance of 2 cm and 86% and 83% for the sample-to-detector distance of 16 cm.

Keywords: structural biomarkers; X-ray scattering; extracellular matrix; cancer detection; machine learning



Copyright: © 2024 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

Breast cancer is currently the leading type of new cancer cases and the second leading cause of cancer mortality for women. More than three hundred thousand new cases of breast cancer and more than forty thousand deaths are estimated for the US in 2024 [1]. While the mortality rate has slightly decreased over the last few years [1], the total number is still worryingly large. The decrease in mortality can be attributed to early diagnosis of breast cancer [2], and thus, the development of early-stage biomarkers is paramount. To address this, we propose monitoring the status of extracellular matrix components using X-ray scattering.

The extracellular matrix (ECM) [3,4] is the noncellular part of tissue with components synthesized mainly by cells. It provides mechanical support for cells and plays a significant role in maintaining tissue homeostasis. The ECM is constantly remodeled to adapt to various stresses [5], and this process generally becomes deregulated in the case of disease [6,7]. Specifically, notable changes in the structure of ECM molecules can indicate a tumor's existence and progression and, correspondingly, serve as biomarker origins [8,9].

Since their discovery, X-rays have been widely used in medicine. However, early and current applications, such as fluorography, radioscopy, and mammography, are based on absorption processes, with X-ray scattering not explicitly exploited. Initially, X-ray crystallography was used to determine atom arrangements in crystalline solids. Later, the possibility of crystallizing biological molecules led to the development of biocrystallography, i.e., the revealing of the molecular content of complex structures; see [10,11] for historical overviews. However, the X-ray scattering approach can be applied to biological molecules or molecular complexes without crystallization, and, if there are periodic structures, they will manifest themselves in scattering patterns. In this, the periodicity, d , of the structure will appear as momentum transfer, $q = 2\pi/d$. While ECM components, such as fibrous proteins (collagens, elastins, fibronectins, and laminins) and glycoproteins, are not crystalline per se, their structural periodicity emerges during synthesis. To separate the studies of such objects from biocrystallography, we introduce the term vitacrystallography, which appears in the title of this paper.

Increasing numbers of reports use X-ray scattering to examine the physicochemical features of breast tissue ECM. Molecular imaging from X-ray scattering has been employed to examine collagen types derived in vivo from intralobular, functionally anomalous fibroblasts associated with invasive tumors. Initially, studies examined tertiary-level collagen structures, leading to new breast tumor microenvironment models [12,13]. Subsequently, multiple studies have examined X-ray scattering (coherent and incoherent) to probe the atomic scales of breast tissues. These have demonstrated high levels of imaging contrast based upon tissue type [14–16]. More recently, the physicochemical properties of ECM microcalcifications (small calcium phosphate deposits typically 10–100s μm in size) frequently associated with breast cancers, especially ductal carcinomas, have been reported. The chemical and microstructural characteristics of apatite (type II) microcalcifications act as immortalized biomarkers of the tissue environment when calcifications form, i.e., at the onset of pathological change. Thus, such features may provide a record of tissue changes from genesis to invasion. Atomic-scale X-ray scattering unlocks this information to indicate both intracellular and extracellular ionic concentrations [17–21] and other characteristics. This is relevant to, for example, elevated intracellular Na and Mg associated with mitogenesis. Such data would impact upon research into overtreatments as Mg is critical to cellular temporal benchmarking.

Another component that can be monitored in relation to cancer progression is lipids. The relationship between cancer and lipid metabolism was first reported in the mid-60s [22,23], with recent studies revisiting this relationship [24–28]. Cancer cells modify lipid metabolism, activating, desaturating, or elongating fatty acids. Moreover, de novo synthesized lipids can differ from those in circulation, affecting the lipid composition of tumors. These changes in lipid profile can be used to establish biomarkers. Currently, the primary analytic tool is mass spectroscopy [24,29], but lipids can be addressed directly using X-ray diffraction. For example, at least two diffraction maxima from breast tissue scattering are uniquely associated with lipid structures. One of them, the wide-angle maximum at $q = 13.9 \text{ nm}^{-1}$, corresponds to inter-fatty-acid molecular distances. It was shown [30–32] that for cancerous tissues, the intensity of this maximum reduces (often to $<3\sigma_{\text{background}}$), while another peak at $q = 20.2 \text{ nm}^{-1}$ concomitantly increases in intensity. This maximum is attributed to the oxygen–oxygen distance in the tetrahedral structure of water [31]. Accordingly, it can be used as a structural biomarker. It should be noted that the values of q are off by 4π in [14,15,33], where it was reported that healthy adipose tissue exhibits a narrow peak at $q = 1.1 \text{ nm}^{-1}$, while for cancerous tissues, this peak appears

shifted to $q = 1.5 \text{ nm}^{-1}$. When multiplied by 4π , it would give values similar to [30–32]. In small-angle scattering, the peak at $q = 1.5 \text{ nm}^{-1}$ has been attributed to the regular packing of triglyceride molecules and is a characteristic of healthy tissue [31,34,35]; moreover, it was shown [35] that this peak is absent in benign tissues. Another work [36] reported a peak at $q = 1.725 \text{ nm}^{-1}$ that appears primarily associated with malignant lesions.

The metastatic spread of breast cancer often involves bone as a secondary site. Although ‘bone homing’ mechanisms are poorly understood, some early indications of the role of local mineral environments are starting to emerge [37]. Many fundamental questions are associated with osteo-tumors that currently hinder the progression of early-stage diagnostics and effective therapies. For example, dormancy processes and mesenchymal–epithelial transitions have been difficult to characterize but are essential for understanding the natural history of metastatic cancers [38,39]. Dormancy and other components (progression) of the metastatic cascade are undoubtedly influenced by tumor microenvironments [40]. X-ray scatter imaging can critically provide information simultaneously on bone mass, architecture, and material ‘quality’ associated with microenvironments [41]. Further, detection of bone cancer at significantly earlier stages than currently possible with conventional radiology within animal models would be possible due to the enhanced contrast afforded by the technique. X-ray scattering has already been demonstrated to be capable of providing new biomarkers for osteoporosis and fracture prediction using atomic scale features (0.4–0.1 nm) averaged over $\sim\text{mm}^3$ mineral volumes [42,43]. Simultaneous architectural imaging can also be employed, using finite element modeling, to provide mechanical properties of the tissue and thus indicate the degradation of bone strength as cancer progresses.

In this paper, we report the analysis of data produced by measurements of X-ray scattering of breast tissues. We use two distinct sample-to-detector distances to address different q -ranges. The distance of 2 cm provides a q -range of up to 22.5 nm^{-1} , while the distance of 16 cm magnifies small-angle scattering with a q -range of up to 4.5 nm^{-1} . These two distances are chosen to examine both ranges associated with cancer-induced lipid transformations. We measure samples from 107 unique patients, obtaining 2958 distinctive scattering patterns. After azimuthal integration, these patterns are converted to dependencies of the scattered intensity on q . The averaged curves exhibit the previously obtained results, with a $q = 1.5 \text{ nm}^{-1}$ peak predominantly appearing in healthy tissues; suppression of the $q = 13.9 \text{ nm}^{-1}$ peak; and the appearance of a $q = 20.2 \text{ nm}^{-1}$ peak in cancerous samples. However, the total number of samples in our research significantly exceeds that of previous studies. It became possible for two reasons. First, many samples can be readily processed as we use a developed table-top diffractometer instead of a synchrotron. Second, a machine-learning technique has been employed for data analysis.

2. Materials and Methods

2.1. Experimental Design

2.1.1. Breast Tissue Specimens

A total of 214 fresh-frozen core biopsy breast tissue specimens totaling 107 patients were utilized for this study and were obtained via the Breast Cancer Now Tissue Biobank. Two specimens per patient were provided for the study to assess heterogeneity. The donors were all female and ranged from 19 to 93 in age and included 60 patients diagnosed with cancer and 47 non-cancer patients. Ethical approval for the project was obtained locally via Keele University (NS-210096), and for the collection and use of specimens via the BCN tissue biobank (NRES Approval Number 21/EE/0072).

2.1.2. Sample Preparation

Each tissue specimen was placed into a bespoke aluminum sample holder of 2 mm thickness utilizing a SPEX™ 6 μm thick mylar window film to seal and secure the tissue within a 5 mm aperture, as shown in Figure 1a. Crucially, the sample holders were designed to ensure the tissues remained hydrated throughout the analysis, which was validated by utilizing porcine tissue prior to scanning human tissue. To ensure consistency

in sample-to-detector distances and measurements, the volume of tissue fully filled the sample holder. Silver behenate powder (Thermoscientific® 045494.06, Waltham, MA, USA) was also scanned (utilizing the bespoke aluminum holders) to allow for accurate sample-to-detector distance measurements.

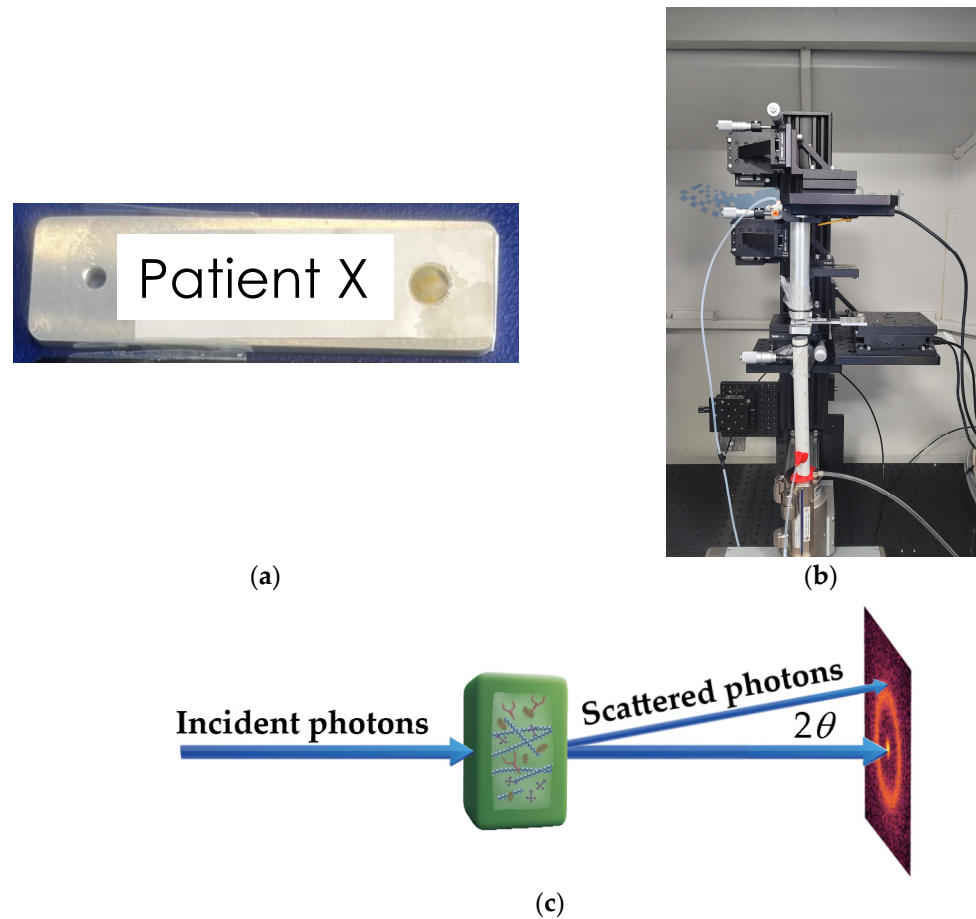


Figure 1. (a) Human breast tissue sealed within a 5 mm aperture bespoke sample holder; (b) EosDx X-ray diffractometer; (c) schematics of the experiment.

2.1.3. X-ray Diffraction (XRD) Measurements

XRD analysis was carried out using a bespoke X-ray diffractometer (Figure 1b) engineered and built by EosDx, Inc. (Menlo Park, CA, USA), a US-based company developing X-ray scattering for medical diagnostics. The radiation produced by the Incoatec Microfocus Source copper source was collected with two multilayer curved mirror optics, resulting in a low-divergence monochromatic beam with a radiation wavelength $\lambda = 0.154$ nm. The two-dimensional detector was an MiniPix SN1442 Si 500 μm detector (ADVACAM, Prague, Czech Republic) with a 256 by 256 pixel array and a 55 by 55 μm pixel size. Schematics of the experiment are shown in Figure 1c. The incident X-ray beam hits the sample containing various biomolecules and is scattered by the angle 2θ . The experimental data were stored as a 256 by 256 matrix of integers representing the photon counts. All the experiments were performed at room temperature (19 °C) under atmospheric pressure. To assess specimen heterogeneity, specimens were mapped with data collected from 10 individual spots on the sample, providing 10 diffractograms per specimen. Each diffractogram was collected for 3 min with a count time of 0.1 s. Background data were collected from an empty aluminum sample holder sealed utilizing mylar (as described above) and subtracted from sample data.

2.2. Data Analysis

The first step of data analysis was removing images from the database that were either empty or lacked a defined beam center, ensuring that only relevant data proceeded to the analysis stage. Following data cleaning, the center of each measured pattern was identified using the Pyfai module, and the precise sample-to-detector distance was determined using a calibration based on AgBH diffraction. The next step was azimuthal integration around the center to obtain radial profiles, effectively translating 2D image data into 1D dependence of the intensity of scattered X-rays on the distance from the center. Initially expressed in pixel units, the radial profiles were converted into momentum transfer $q = (4\pi \sin \theta) / \lambda$, where $\sin 2\theta$ was calculated as the ratio of the distances from the pixel to the center and from the sample to the detector; see Figure 1c.

The dataset was clustered into two groups corresponding to the two different experimental sample-to-detector distances. The q -ranges were interpolated within each cluster to achieve a uniform length of 50 points, enhancing comparability, with a 4.5% cut from both sides of the q -range. This step involved the removal of outliers using the z-score technique to ensure the reliability of the interpolation process.

Prior to classification, the radial profiles underwent normalization and scaling. This step, executed with the scikit-learn library [44], involved adjusting the data within each cluster to have a standard distribution. The processed data from each cluster underwent a randomized separation (200 times) with respect to patients, with a train-test split of an approximately 60/40 ratio, forming not-overlapping subsets, i.e., training was performed using certain patients, and the test was accomplished with the other subset.

The machine learning algorithm from the scikit-learn library, Random Forest Classifier, was employed to classify the samples into cancerous and non-cancerous categories. This algorithm was chosen for its robustness and ability to handle complex patterns in data. The training phase involved adjusting the models to best fit the training data, while the testing phase assessed their performance in classifying unseen data. The number of trees was set to 100; the depth of trees was set to 10; and the learning rate of 0.01 was set for Random Forest. The learning was performed for all train-test subsets; 200 for each cluster.

We evaluated our model using sensitivity and specificity as our performance metrics. Sensitivity refers to the proportion of tumor samples that were accurately identified, while specificity relates to the accuracy in identifying control samples. To determine the most effective threshold for classification, we selected the one that brought our model's performance closest to that of an ideal classifier, as indicated on the receiver operating characteristic (ROC) curve. The area under the ROC curve (AUC) is another critical measure, with values closer to one indicating superior performance.

3. Results

The samples were measured at two sample-to-detector distances. The descriptions of the collected datasets are shown in Table 1.

Table 1. Datasets for X-ray scattering data measured at two different distances of 2 and 16 cm.

Clusters	Diagnosis		Total
	Cancer	Non-Cancer	
Cluster 2 cm	825	1463	2288
Cluster 16 cm	306	364	670
	1131	1827	2958

Typical results of the measurements are shown in Figure 2.

As shown in Figure 2, the scattering patterns are almost isotropic, and the azimuthal integration can be performed without losing information. As a result of this integration, the scattered intensity can be represented as a function of the distance to the center, which,

in turn, can be rearranged in terms of the transfer momentum magnitude. The obtained dependencies of the intensity of the scattered X-rays on the transfer momentum are shown in Figure 3. The 2 cm cluster includes the region of $2.3\text{--}22.5\text{ nm}^{-1}$ of the transfer momentum magnitude, while the 16 cm cluster covers the $0.44\text{--}4.29\text{ nm}^{-1}$ range.

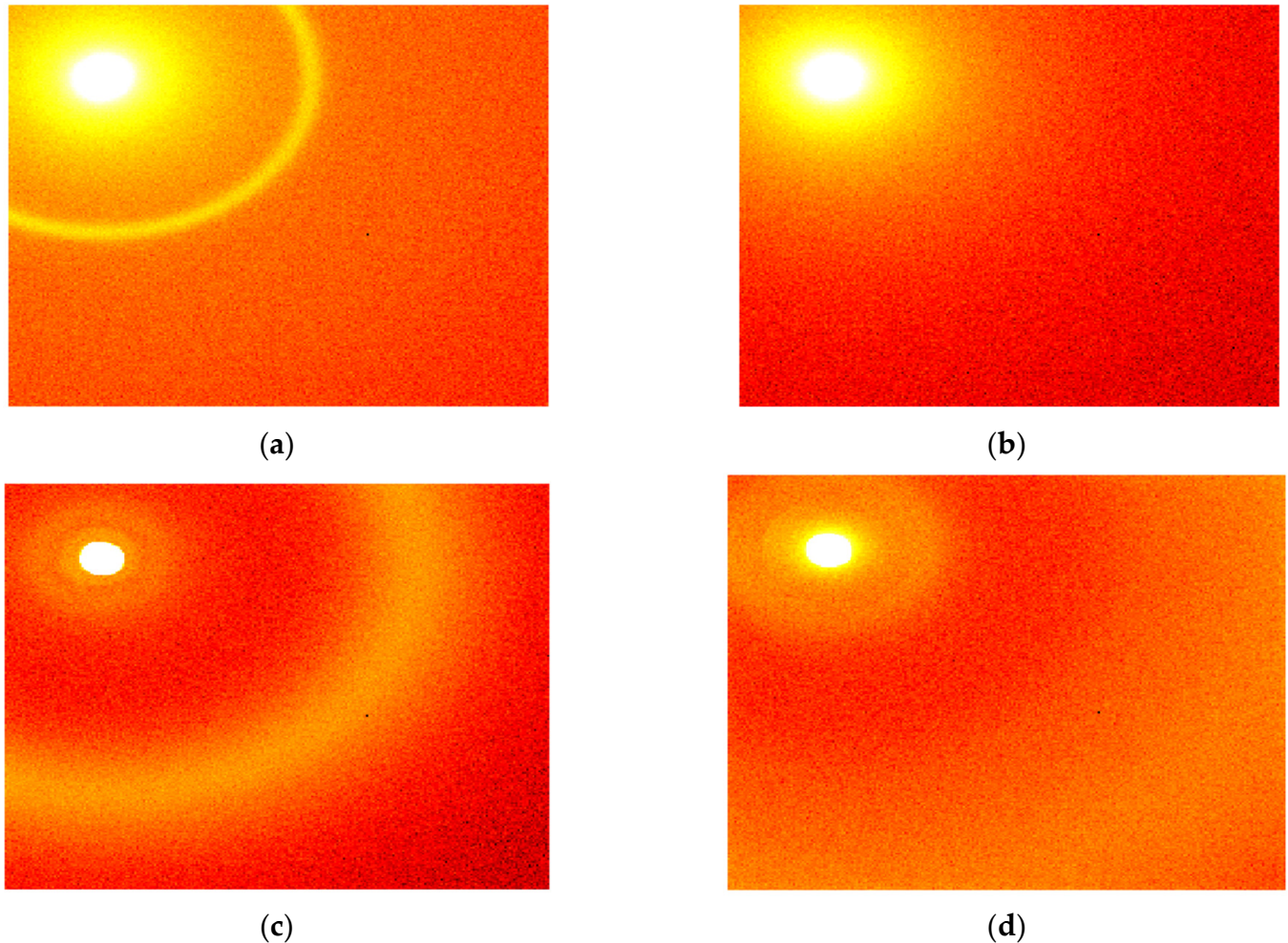


Figure 2. Typical X-ray scattering patterns obtained for (a) non-cancer samples measured at the sample-to-detector distance of 16 cm, (b) cancer samples measured at 16 cm, (c) non-cancer samples measured at 2 cm, and (d) cancer samples measured at 2 cm.

If we compare the averaged intensities, shown as black dashed curves, we can confirm the results of previous studies [37–42] related to lipid metabolism. For the 16 cm measurements, the peak at $q = 1.5\text{ nm}^{-1}$ is very pronounced in non-cancer samples (Figure 3a) but much less visible in cancer samples (Figure 3b). For the 2 cm measurements, the peak at $q = 14\text{ nm}^{-1}$ is clearly seen in non-cancer samples (Figure 3c) but less pronounced than the peak at $q = 20\text{ nm}^{-1}$ in cancer samples (Figure 3d). These findings confirm previously obtained results relating these peak modifications to lipid metabolism [30–32,34,35]. However, the individual measurements can be significantly different from these general trends, as seen in all the panels of Figure 3. In particular, the 1.5 nm^{-1} peak appears in many cancer samples (Figure 3b). In contrast, the 20 nm^{-1} peak dominates the one at 14 nm^{-1} for many non-cancer samples (Figure 3c). Correspondingly, to reveal all the tendencies and facilitate the applications of this approach to actual diagnostics, we applied machine learning techniques.

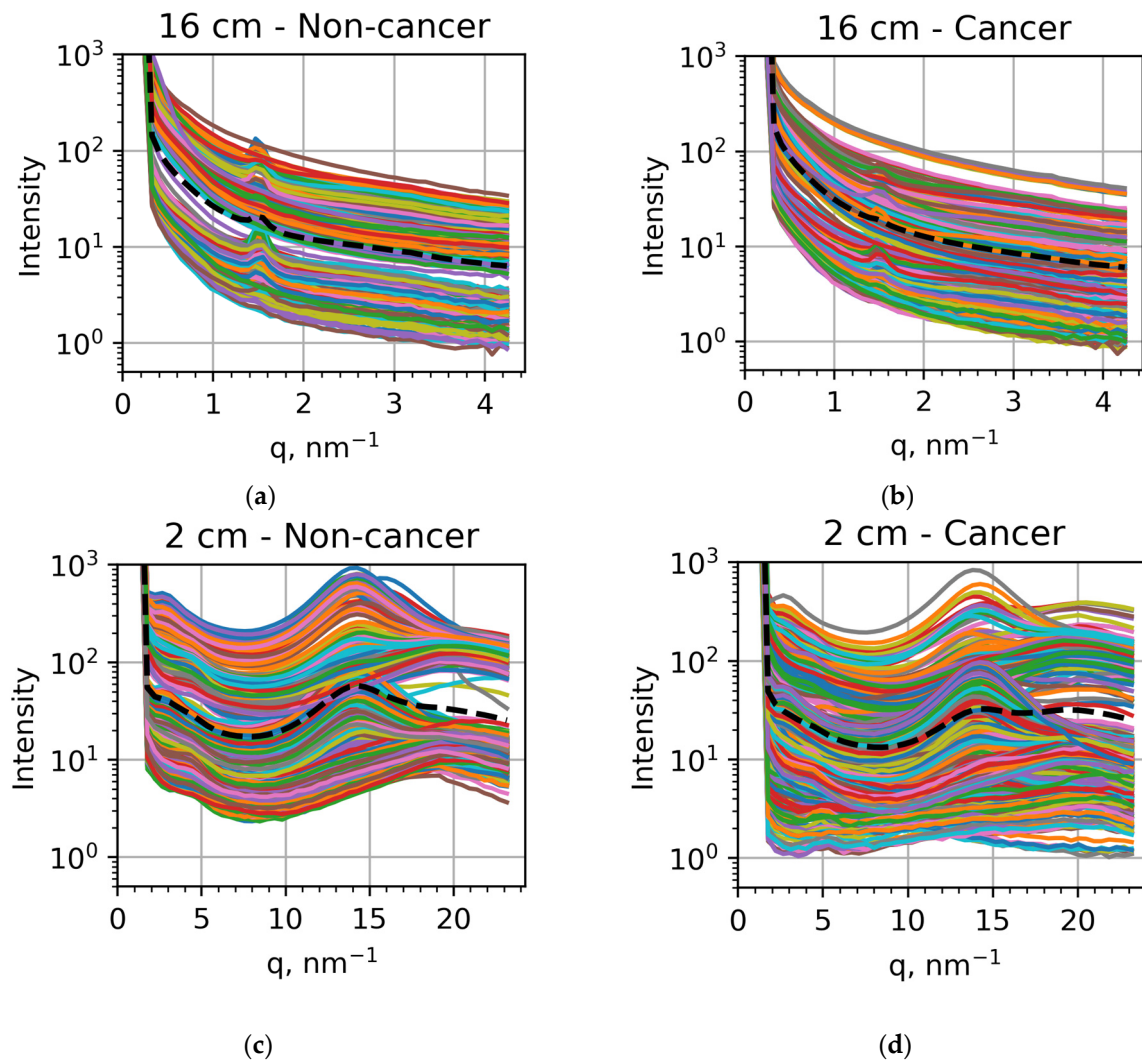


Figure 3. Dependencies of the scattered intensities on the transfer momentum obtained for (a) non-cancer samples measured at the sample-to-detector distance of 16 cm, (b) cancer samples measured at 16 cm, (c) non-cancer samples measured at 2 cm, and (d) cancer samples measured at 2 cm. The intensities for each specific sample are showed at different colors, with the averaged intensities shown as black dashed curves.

We trained the Random Forest Classifier procedure on 60% of measurements, using the remaining 40% for the test to determine the outcome metrics, with the same percentage of cancer patients in both clusters as in the total dataset. Several measurements of a single patient were kept in the same training or test clusters. Such training/test sequences were performed 200 times. The ROC curves for the test sets are shown in Figure 4 for both 16 cm and 2 cm measurements.

With each of the 200 realizations exhibited in grey, the best and the worst are shown in red and blue, respectively. The average curve is black. The corresponding performance metrics are presented in Table 2.

Table 2. Random Forest performance metrics for 2 cm and 16 cm datasets.

Metrics Distances (cm)	AUC ROC		Sensitivity		Specificity	
	2	16	2	16	2	16
Random Forests (min)	64	58	56	78	73	36
Random Forests (max)	92	89	80	86	92	83
Random Forests (average)	78	73.5	68	82	82.5	59.5

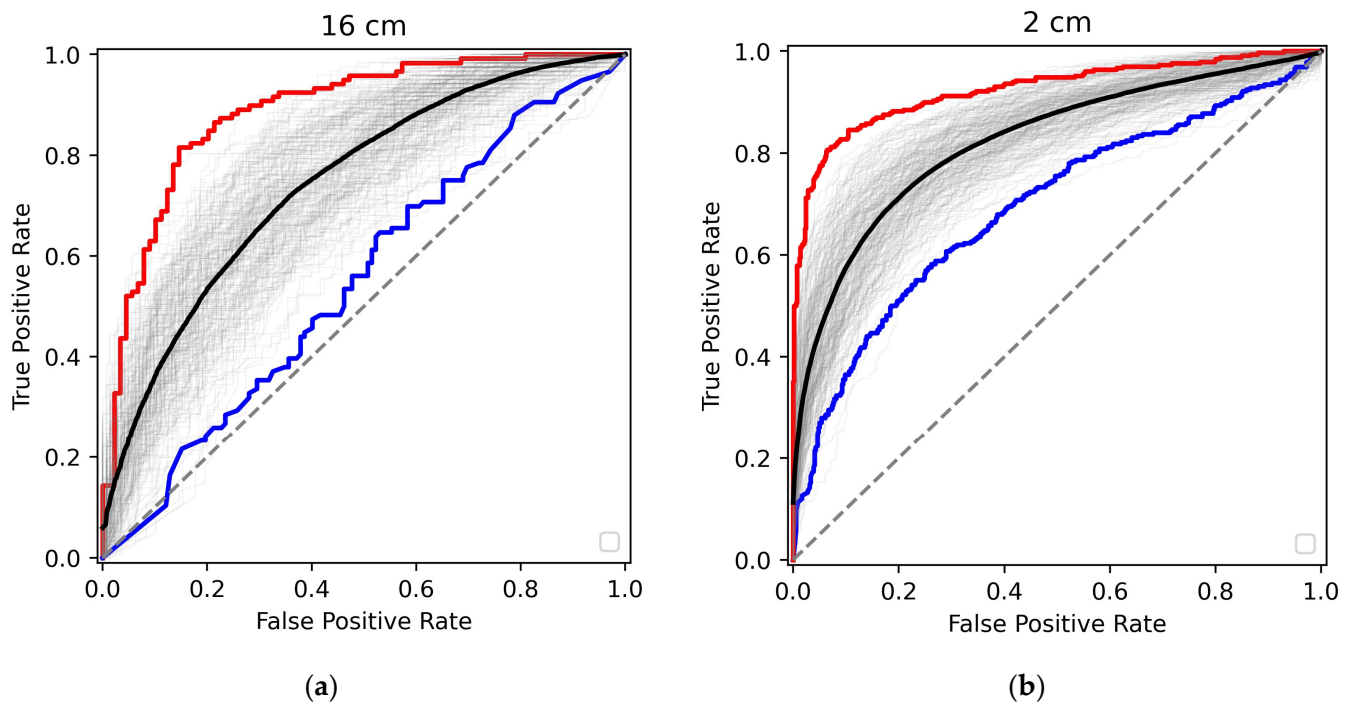


Figure 4. The ROC curves for various realizations of the training-test shuffles. (a) Measurements at the sample-to-detector distance of 16 cm and (b) at the sample-to-detector distance of 2 cm. The best (highest ROC AUC), worst (lowest ROC AUC), and averaged realizations are shown in red, blue, and black, respectively.

4. Discussion

It is evident from Figure 4 and Table 2 that the Random Forrest Classifier approach allows promising cancer/non-cancer binary diagnostics. For the best realization, the AUC ROC reached 94% for the 2 cm measurements and 87% for the 16 cm measurements. The metrics for the average realization are slightly worse but still very good. As illustrated in Table 2, the test specificity is demonstrated to possess credible metrics consistent with confirmatory diagnostic tests following initial positive screening. In such cases, high specificity becomes critical to ensure that patients identified as having cancer truly have the disease and to reduce the number of false positives and overtreatment. In general, the performance metrics are better for the 2 cm experiments. We believe this is due to the extremely small q -range, with significant contributions from the primary beam and small-angle amorphous scattering, which have more minor effects on data analysis in this case.

It should be emphasized that the most prominent features for both visual analyses based on Figure 3 and performance metrics obtained in machine learning are related to lipid metabolism. In the presence of a tumor, regular triglyceride packing is broken, suppressing the 1.5 nm^{-1} peak. Tumor cells also affect fatty acids, so the 13.9 nm^{-1} peak, corresponding to inter-fatty-acid molecular distances in healthy tissues, becomes less pronounced. It is crucial to include the other components of ECM in our analysis. However, the resolution

on our diffractometer was too low to observe discrete collagen peaks at a low q , which is the direction of future improvements.

In the present paper, the measurements were performed on small samples obtained from biopsies. However, the same structural biomarkers determined in our studies would also appear in large samples and, eventually, in *in vivo* procedures. The developed tabletop diffractometer should be modified for these future applications. Instead of the current Cu source, with a penetration depth of several mm, a Mo or Ag source will be used. They have much shorter wavelengths, leading to more negligible absorption and a penetration depth of several hundred mm. Moreover, the design of the diffractometer must be changed to accommodate *in vivo* measurements, and the current device is a prototype on the path to this ultimate goal. As biomarker determination is based on X-ray diffraction, not absorption, the received dose will be negligible. It should also be emphasized that at least small-sample studies are not limited to breast cancer. X-ray scattering measurements can determine structural biomarkers associated with other types of cancer and even other diseases.

5. Conclusions

In conclusion, we performed X-ray scattering measurements of breast tissue samples from 107 patients at two sample-to-detector distances of 2 and 16 cm. After azimuthal integration of the acquired patterns, we obtained the dependencies of scattered intensity on the momentum transfer magnitudes, exhibiting lipid structure features. The peaks associated with the regular packing of triglyceride molecules (at 1.5 nm^{-1}) and with inter-fatty-acid molecular distances (at approximately 14 nm^{-1}) were observed predominantly in non-cancer tissues, which confirms the results of previous studies, on average. To examine the feasibility of cancer detection on an individual level, we performed a statistical analysis based on the machine-learning Random Forest Classifier approach. We separated the total dataset into 60/40 training/test splits and repeated this separation 200 times, with the training and test splits not overlapping regarding patients. The obtained performance metrics vary from very good to excellent, with the sensitivity and specificity reaching 86% and 83%, respectively, for the best realization. We believe such results show great promise for future non-invasive technology for breast cancer detection.

Author Contributions: Conceptualization, L.M. and P.L.; methodology, S.D., C.G., K.R. and P.L.; software, S.D., B.B. and J.F.; validation, L.M. and P.L.; investigation, B.G., A.A., L.A. and C.G.; data curation, S.D., B.B., J.F. and C.G.; writing—original draft preparation, S.D., C.G. and L.M.; writing—review and editing, K.R.; visualization, S.D.; supervision, C.G. and K.R.; project administration, C.G. and P.L.; funding acquisition, P.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Tissue was collected under NRES Approval number 21/EE/0072, with university ethics approval granted for this research (NS-210096).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study by the Breast Cancer Now Biobank.

Data Availability Statement: All produced data, including raw measurement data and the Jupiter notebooks for statistical analysis, will be made readily available by authors on request.

Acknowledgments: The authors wish to acknowledge the roles of the Breast Cancer Now Tissue Bank in collecting and making available the samples and data, and the patients who have generously donated their tissues and shared their data to be used in the generation of this publication.

Conflicts of Interest: S.D., B.B. and P.L. are employees of Matur UK, Ltd., UK. B.B., J.F. and P.L. are employees of EosDx, Inc., USA. C.G., K.R. and L.M. are consultants of EosDx, Inc., USA. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Siegel, R.L.; Giaquinto, A.N.; Jemal, A. Cancer statistics, 2024. *CA Cancer J. Clin.* **2024**, *74*, 12–49. [[CrossRef](#)]
2. Taylor, C.; McGale, P.; Probert, J.; Broggio, J.; Charman, J.; Darby, S.C.; Kerr, A.J.; Whelan, T.; Cutter, D.J.; Mannu, G.; et al. Breast cancer mortality in 500,000 women with early invasive breast cancer diagnosed in England, 1993–2015: Population based observational cohort study. *BMJ* **2023**, *381*, e074684. [[CrossRef](#)] [[PubMed](#)]
3. Frantz, C.; Stewart, K.M.; Weaver, V.M. The extracellular matrix at a glance. *J. Cell Sci.* **2010**, *123*, 4195–4200. [[CrossRef](#)]
4. Mecham, R.P. Overview of extracellular matrix. *Curr. Protoc. Cell Biol.* **2012**, *10*, 10.1.1–10.1.16. [[CrossRef](#)] [[PubMed](#)]
5. Hynes, R.O. Extracellular matrix: Not just pretty fibrils. *Science* **2009**, *326*, 1216–1219. [[CrossRef](#)] [[PubMed](#)]
6. Theocharis, A.D.; Manou, D.; Karamanos, N.K. The extracellular matrix as a multitasking player in disease. *FEBS J.* **2019**, *286*, 2830–2869. [[CrossRef](#)] [[PubMed](#)]
7. Sala, M.; Ros, M.; Saltel, F. A Complex and Evolutive Character: Two Face Aspects of ECM in Tumor Progression. *Front. Oncol.* **2020**, *10*, 1620. [[CrossRef](#)]
8. Insua-Rodríguez, J.; Oskarsson, T. The extracellular matrix in breast cancer. *Adv. Drug Del. Rev.* **2017**, *97*, 41–55. [[CrossRef](#)]
9. Petersen, E.V.; Chudakova, D.A.; Skorova, E.Y.; Anikin, V.; Reshetov, I.V.; Mynbaev, O.A. The Extracellular Matrix-Derived Biomarkers for Diagnosis, Prognosis, and Personalized Therapy of Malignant Tumors. *Front. Oncol.* **2020**, *10*, 575569. [[CrossRef](#)]
10. Shi, Y.A. Glimpse of Structural Biology through X-Ray Crystallography. *Cell* **2014**, *159*, 995–1014. [[CrossRef](#)]
11. Higgins, M.K.; Lea, S.M. On the state of crystallography at the dawn of the electron microscopy revolution. *Curr. Op. Struct. Biol.* **2017**, *46*, 95–101. [[CrossRef](#)] [[PubMed](#)]
12. Lewis, R.A.; Rogers, K.D.; Hall, C.J.; Towns-Andrews, E.; Slawson, S.; Evans, A.; Pinder, S.E.; Ellis, I.O.; Boggis, C.R.M.; Hufton, A.P.; et al. Breast cancer diagnosis using scattered X-rays. *J. Synchrotron Rad.* **2000**, *7*, 348–352. [[CrossRef](#)] [[PubMed](#)]
13. Falzon, G.; Pearson, S.; Murison, R.; Hall, C.; Siu, K.; Evans, A.; Rogers, K.; Lewis, R. Wavelet-based feature extraction applied to small-angle x-ray scattering patterns from breast tissue: A tool for differentiating between tissue types. *Phys. Med. Biol.* **2006**, *51*, 2465. [[CrossRef](#)] [[PubMed](#)]
14. Kidane, G.; Speller, R.D.; Royle, G.J.; Hanby, A.M. X-ray scatter signatures for normal and neoplastic breast tissues. *Phys. Med. Biol.* **1999**, *44*, 1791. [[CrossRef](#)] [[PubMed](#)]
15. Poletti, M.E.; Gonçalves, O.D.; Mazzaro, I. Coherent and incoherent scattering of 17.44 and 6.93 keV X-ray photons scattered from biological and biological-equivalent samples: Characterization of tissues. *X-ray Spectrom.* **2002**, *31*, 57–61. [[CrossRef](#)]
16. Cunha, D.M.; Oliveira, O.R.; Pérez, C.A.; Poletti, M.E. X-ray scattering profiles of some normal and malignant human breast tissues. *X-ray Spectrom.* **2006**, *35*, 370–374. [[CrossRef](#)]
17. Kunitake, J.A.M.; Choi, S.; Nguyen, K.X.; Lee, M.M.; He, F.; Sudilovsky, D.; Morris, P.G.; Jochelson, M.S.; Hudis, C.A.; Muller, D.A.; et al. Correlative imaging reveals physicochemical heterogeneity of microcalcifications in human breast carcinomas. *J. Struct. Biol.* **2017**, *202*, 25–34. [[CrossRef](#)]
18. Sharma, T.; Radosevich, J.A.; Pachori, G.; Mandal, C.C. A Molecular View of Pathological Microcalcification in Breast Cancer. *J. Mammary Gland. Biol. Neoplasia* **2016**, *21*, 25–40. [[CrossRef](#)] [[PubMed](#)]
19. Hajibeigi, A.; Nasr, K.; Udayakumar, D.; Nham, K.; Lenkinski, R.E. Breast Tumor Microcalcification Induced by Bone Morphogenetic Protein-2: A New Murine Model for Human Breast Tumor Diagnosis. *Contr. Med. Mol. Imag.* **2018**, *2018*, 2082154. [[CrossRef](#)]
20. Bonfiglio, R.; Scimeca, M.; Toschi, N.; Pistolese, C.A.; Giannini, E.; Antonacci, C.; Ciuffa, S.; Tancredi, V.; Tarantino, U.; Albonici, L.; et al. Radiological, Histological and Chemical Analysis of Breast Microcalcifications: Diagnostic Value and Biological Significance. *J. Mammary Gland. Biol. Neoplasia* **2018**, *23*, 89–99. [[CrossRef](#)]
21. Bonfiglio, R.; Scimeca, M.; Urbano, N.; Bonanno, E. Breast microcalcifications: Biological and diagnostic perspectives. *Future Oncol.* **2018**, *14*, 3097–3099. [[CrossRef](#)]
22. Spector, A.A.; Steinberg, D. Relationship between fatty acid and glucose utilization in Ehrlich ascites tumor cells. *J. Lipid Res.* **1966**, *7*, 657–663. [[CrossRef](#)] [[PubMed](#)]
23. Tator, C.H.; Evans, J.R.; Olszewski, J. Tracers for the detection of brain tumors. Evaluation of radioiodinated human serum albumin and radioiodinated fatty acid. *Neurology* **1966**, *16*, 650–661. [[CrossRef](#)] [[PubMed](#)]
24. Butler, L.M.; Perone, Y.; Dehairs, J.; Lupien, L.E.; de Laat, V.; Talebi, A.; Loda, M.; Kinlaw, W.B.; Swinnen, J.V. Lipids and cancer: Emerging roles in pathogenesis, diagnosis and therapeutic intervention. *Adv. Drug Del. Rev.* **2020**, *159*, 245–293.
25. Duong, L.K.; Corbali, H.I.; Riad, T.S.; Ganjoo, S.; Nanez, S.; Voss, T.; Barsomium, H.; Welsh, J.; Cortez, M.A. Lipid metabolism in tumor immunology and immunotherapy. *Front. Oncol.* **2023**, *13*, 118727. [[CrossRef](#)] [[PubMed](#)]
26. Jin, H.-R.; Wang, J.; Wang, Z.-J.; Xi, M.-J.; Xia, B.-H.; Deng, K.; Yang, J.-L. Lipid metabolic reprogramming in tumor microenvironment: From mechanisms to therapeutics. *J. Hem. Oncol.* **2023**, *16*, 103. [[CrossRef](#)] [[PubMed](#)]
27. Goswami, S.; Zhang, Q.; Celik, C.E.; Reich, E.M.; Yilmaz, O.H. Dietary fat and lipid metabolism in the tumor microenvironment. *Biochim. Biophys. Acta Rev. Cancer* **2023**, *1878*, 188984. [[CrossRef](#)] [[PubMed](#)]
28. Zhao, Y.; Zhang, X.; An, M.; Zhang, J.; Liu, Y. Recent advancements in nanomedicine based lipid metabolism for tumour immunotherapy. *J. Drug Target.* **2023**, *31*, 1050–1064. [[CrossRef](#)]
29. Bandu, R.; Mok, H.J.; Kim, K.P. Phospholipids as cancer biomarkers: Mass spectrometry-based analysis. *Mass Spectrom. Rev.* **2018**, *37*, 107–138. [[CrossRef](#)]

30. Conceicao, A.L.C.; Antoniassi, M.; Poletti, M.E. Assessment of the differential linear coherent scattering coefficient of biological samples. *Nucl. Instr. Meth. Phys. Res. A* **2010**, *619*, 67–70. [[CrossRef](#)]
31. Conceicao, A.L.C.; Meehan, K.; Antoniassi, M.; Piacenti-Silva, M.; Poletti, M.E. The influence of hydration on the architectural rearrangement of normal and neoplastic human breast tissues. *Heliyon* **2019**, *5*, e01219. [[CrossRef](#)] [[PubMed](#)]
32. Friedman, J.; Blinchevsky, B.; Slight, M.; Tanaka, A.; Lazarev, A.; Zhang, W.; Aram, B.; Ghadimi, M.; Lomis, T.; Mourokh, L.; et al. Structural Biomarkers for Breast Cancer Determined by X-Ray Diffraction. In *Quantum Effects and Measurement Techniques in Biology and Biophotonics*; Aiello, C., Polyakov, S.V., Derr, P., Eds.; SPIE: San Francisco, CA, USA, 2024; Volume 12863, p. 1286302.
33. Moss, R.M.; Amin, A.S.; Crews, C.; Purdie, C.A.; Jordan, L.B.; Iacoviello, F.; Evans, A.; Speller, R.D.; Vinnicombe, S.J. Correlation of X-ray diffraction signatures of breast tissue and their histopathological classification. *Sci. Rep.* **2017**, *7*, 12998. [[CrossRef](#)] [[PubMed](#)]
34. Suortti, P.; Fernandez, M.; Urban, V. Comments on Synchrotron fibre diffraction identifies and locates foetal collagenous breast tissue associated with breast carcinoma by V. J. James (2002). *J. Synchrotron Rad.* **2003**, *10*, 198. [[CrossRef](#)]
35. Mohd Sobri, S.N.; Abdul Sani, S.F.; Sabtu, S.N.; Looi, L.M.; Chiew, S.F.; Pathmanathan, D.; Chio-Srichan, S.; Bradley, D.A. Structural Studies of Epithelial Mesenchymal Transition Breast Tissues. *Sci. Rep.* **2020**, *10*, 1997. [[CrossRef](#)] [[PubMed](#)]
36. Arboleda, C.; Lutz-Bueno, V.; Wang, Z.; Villanueva-Perez, P.; Guizar-Sicairos, M.; Liebi, M.; Varga, Z.; Stampanoni, M. Assessing lesion malignancy by scanning small-angle X-ray scattering of breast tissue with microcalcifications. *Phys. Med. Biol.* **2019**, *64*, 155010. [[CrossRef](#)] [[PubMed](#)]
37. He, F.; Chiou, A.E.; Loh, H.C.; Lynch, M.; Seo, B.R.; Song, Y.H.; Lee, M.J.; Hoerth, R.; Bortel, E.L.; Willie, B.M.; et al. Multiscale characterization of the mineral phase at skeletal sites of breast cancer metastasis. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 10542–10547. [[CrossRef](#)] [[PubMed](#)]
38. Sosa, M.S.; Bragado, P.; Aguirre-Ghiso, J.A. Mechanisms of disseminated cancer cell dormancy: An awakening field. *Nat. Rev. Cancer* **2014**, *14*, 611–622. [[CrossRef](#)] [[PubMed](#)]
39. Esposito, M.; Mondal, N.; Greco, T.; Wei, Y.; Spadazzi, C.; Lin, S.-C.; Zheng, H.; Cheung, C.; Magnani, J.; Lin, S.-H.; et al. Bone Vascular Niche E-selectin Induces Mesenchymal-Epithelial Transition and Wnt Activation in Cancer Cells to Promote Bone Metastasis. *Nat. Cell Biol.* **2019**, *21*, 627–639. [[CrossRef](#)]
40. Kan, C.; Vargas, G.; Le Pape, F.; Clézardin, P. Cancer Cell Colonisation in the Bone Microenvironment. *Int. J. Mol. Sci.* **2016**, *17*, 1674. [[CrossRef](#)]
41. Zanghellinia, B.; Grünewald, T.A.; Burghammer, M.; Rennhofer, H.; Liegl-Atzwanger, B.; Leithner, A.; Lichtenegger, H.C. High-resolution large-area imaging of nanoscale structure and mineralization of a sclerosing osteosarcoma in human bone. *J. Struct. Biol.* **2019**, *207*, 56–66. [[CrossRef](#)]
42. Greenwood, C.; Clement, J.; Dicken, A.; Evans, J.P.O.; Lyburn, I.; Martin, R.M.; Rogers, K.; Stone, N.; Zioupos, P. Towards new material biomarkers for fracture risk. *Bone* **2016**, *93*, 55–63. [[CrossRef](#)] [[PubMed](#)]
43. Dicken, A.J.; Evans, J.P.O.; Rogers, K.D.; Stone, N.; Greenwood, C.; Godber, S.X.; Clement, J.G.; Lyburn, I.D.; Martin, R.M.; Zioupos, P. Classification of fracture and non-fracture groups by analysis of coherent X-ray scatter. *Sci. Rep.* **2016**, *6*, 29011. [[CrossRef](#)] [[PubMed](#)]
44. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.