Correlation of breast cancer biology and x-ray scatter data

by

Prarthana Pasricha

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science (MSc) in Physics

The Office of Graduate Studies
Laurentian University
Sudbury, Ontario, Canada

THESIS DEFENCE COMMITTEE/COMITÉ DE SOUTENANCE DE THÈSE

Laurentian Université/Université Laurentienne

Office of Graduate Studies/Bureau des études supérieures

Title of Thesis Titre de la thèse	Correlation of breast cancer biology and x-ray scatter data		
Name of Candidate Nom du candidat	Pasricha, Prarthana		
Degree Diplôme	Master of Science		
Department/Program Département/Programme	Physics	Date of Defence Date de la soutenance	August 24, 2022

APPROVED/APPROUVÉ

Thesis Examiners/Examinateurs de thèse:

Dr. Robert Leclair / Dr. Dean Millar
(Supervisor/Directeur(trice) de thèse)

Dr. Aseem Kumar
(Committee member/Membre du comité)

Dr. Gennady Chitov
(Committee member/Membre du comité)

Approved for the Office of Graduate Studies
Approuvé pour le Bureau des études supérieures
Tammy Eger, PhD
Vice-President Research (Office of Graduate Studies)
Vice-rectrice à la recherche (Bureau des études supérieures
Laurentian University / Université Laurentienne

Dr. Ghirmay Kidane
(External Examiner/Examinateur externe)

Vice-rectrice à la recherche (Bureau des études supérieures)
Laurentian University / Université Laurentienne

ACCESSIBILITY CLAUSE AND PERMISSION TO USE

In memory of my golden labrador Mishy, who we lost to cancer this year. She was the best and the bravest. We miss her beyond words.

Abstract

Wide angle x-ray scatter (WAXS) imaging can potentially be used to diagnose ductal carcinoma in situ (DCIS) in breast biopsies. In this work, WAXS signals are used to model healthy and diseased breast tissue. It is accomplished by analysis of the breast tissue's differential linear scattering coefficients 'μ_s'. The tissue samples are assumed to consist of varying amount of *fat*, *stroma* and *cells* and their μ_s coefficients can be expressed by the following sum: $\mu_s(x) = a_{fat} \cdot \mu_s^{fat}(x) + a_{str} \cdot \mu_s^{str}(x) + a_{cell} \cdot \mu_s^{cell}(x)$, where $x = (1/\lambda) \sin(\theta/2)$ is the momentum transfer argument, λ is the x-ray wavelength and θ is the scatter angle. Data for breast tissues was taken from literature and that for stroma was estimated using histology studies. The fractional volumes of each of the three components (a_{fat}, a_{str}, a_{cell}) were evaluated using singular value decomposition (SVD). Fitting results suggest cancer consists of 44% cells as compared to only 8% and 7% for normal and benign tissue, respectively. These results support the argument that an uncontrolled proliferation of cells is the hallmark of cancer. The fat decreases from 56% (normal) to 26% (benign) and to only 3.4% in cancer. This is justified by the fact that invasive ductal carcinoma cells use fat to fuel their growth and spread. Stromal increase from 35% in normal to 67% in benign suggest that benign diseases such as fibroadenoma are characterized by a surge in fibrous tissue. The model predictions of the compositions for normal, benign and malignant tissue are promising. It gives insight into how cancer interacts with its microenvironment.

While scatter studies on cells and fat have been done in the past, not much is known about stroma in the same context. Studies show that stroma promotes cancer growth and *metastasis*, making it an important component to focus on. The second part of this research works on this aspect. A model for stroma made up of 80% collagen, 10% elastin and 10% versican is introduced. μ_s for stroma is

then determined using Independent Atomic Model (*IAM*) approximation. The scatter signal obtained via this newly introduced stroma model matches with the previous estimation, mainly at the higher values of momentum transfer values. This is because IAM fails at the lower regime due to interference phenomenon. It is observed that the scatter values obtained using IAM are not very sensitive to the change in composition of stroma. In most of the cases, the percentage change with reference to the original model is less than 1%. In the event in which elastin is considered to be the major constituent, a difference of 2.41% to 3.49% is seen. This model gives important information about the structure and function of stroma and highlights its role in cancer progression.

Acknowledgements

As for any piece of research that results in the production of a thesis, on the cover there should not be only the name of the researcher, but also the names of all those who provided guidance, encouragement and assistance, and without whom I would not have completed my thesis. I offer my warmest gratitude to all those people, my heroes, for their time, love and energy.

Heartfelt thanks to my thesis supervisor Dr. Robert Leclair, for his constant guidance and willingness to share his vast knowledge that made me understand the depths of the project and allowed me to grow as an independent researcher. While I am not the best student, he surely is an advisor extraordinaire. Thanks for being patient and for your encouragement, that kept me going. I'd like to thank my committee members Dr. Gennady Chitov and Dr. Aseem Kumar for their brilliant comments and suggestions. Having expertise from two different fields was very beneficial for this interdisciplinary research. Special thanks to Dr. Eduardo Galiano and Dr. Christine Kraus for checking on me time to time and for their advice and influence throughout my master's degree. This work would not have been possible without Dr. Dean Millar, thanks for pushing me to the finish line.

I want to thank my parents and my brother for their continuous support and endless love. Their belief in me has kept my spirits and motivation high during this project. I am extremely thankful for their prayers and sacrifices for educating me and preparing me for future.

I am forever grateful to the friends I made in Sudbury: Mansi, Ashmeet, Falshruti, Jerin, Gurbir. Thanks for being there!

Finally, thankyou God for being there every step of the way ♡

Table of Contents

List of figures

List of tables

Glossary

Amino-acids: An amino acid is an organic molecule that is made up of a basic amino group ($-NH_2$), an acidic carboxyl group ($-COOH$), and an organic *R* group (or side chain) that is unique to each amino acid

Apocrine: A gland and especially a sweat gland that secretes a viscous fluid into a hair follicle (as in the armpit or the groin), is lined with a single layer of usually columnar cells, and typically does not become active until puberty

Apoptosis: Programmed cell death that occurs in multicellular organisms

Atypical: As in atypical cell growth. Cells that are not normal but are not cancerous. Atypical cells could become a cancer over time or may increase a person's risk of cancer

Benign: Refers to a condition, tumor or growth that is not cancerous

Biopsy: An examination of tissue removed from a living body to discover the presence, cause or extent of a disease

Breast tomosynthesis: 3-D mammography, advanced form of breast imaging, or mammography, that uses low dose x-ray system and computer reconstructions to create 3D images of the breast. It aids in early detection and diagnosis of breast disease

Collagen: The main structural protein found in skin and other connective tissue

Connective tissue: Tissue that connects, supports, binds or separates other tissues or organs, typically having relatively few cells embedded in an amorphous matrix, often with collagen or other fibers, and including cartilaginous, fatty, and elastic tissues

Cancer: Cancer is a large group of diseases that can start in almost any organ or tissue of the body when abnormal cells grow uncontrollably, go beyond their usual boundaries to invade parts of the body and/or spread to other organs. The latter process is called metastasizing and is a major cause of death from cancer. A neoplasm and malignant tumour are other common names for cancer

Carcinoma: carcinoma is the most common type of cancer. It begins in the epithelium found in skin or, more commonly, the lining of body organs, for example: breast, prostrate, lung, stomach, or bowel

Cyst: A cyst is a sac that may be filled with air, fluid or other material

Diffraction: Refers to various phenomena that occurs when a wave encounters an obstacle or opening. It is defined as the interference or bending of waves around the corners of an obstacle or through an aperture into the region of geometrical shadow of the obstacle or aperture

Desmoplasia: Also known as desmoplastic reaction is the term used by pathologists to refer to the growth of fibrous tissue around the disease, usually cancer

Elastin: It is a protein that in humans is encoded by the ELN gene. Elastin is a key component of the extracellular matrix. It is highly elastic and present in connective tissues allowing many tissues in the body to resume their shape after stretching or contracting

Extracellular matrix: The extracellular matrix (ECM) is the non-cellular component present within all tissues and organs, and provides not only essential physical scaffolding for the cellular constituents but also initiates crucial biochemical and biomechanical cues that are required for tissue morphogenesis, differentiation and homeostasis

Fibroblast: A fibroblast is a type of cell that contributes to the formation of connective tissue, a fibrous cellular material that supports and connects other tissues or organs in the body. Fibroblasts secrete collagen proteins that help maintain the structural framework of tissues. They also play an important role in healing wounds

Glycine: (Gly) is an amino acid that has a single hydrogen atom as a side chain. It is a constituent of most proteins

Hydroxyproline: (Hyp) is an amino acid occurring in some proteins especially collagen

Lumen: Cavity or channel within a tube or tubular organ

Malignant: A term used to describe cancer. Malignant cells grow in an uncontrolled way and can invade nearby tissues and spread to other parts of the body through the blood and lymph system

Metastasis: The spread of cancer cells from the place where they first formed to another part of the body

Myofibroblast: Myofibroblasts are the mesenchymal cell type responsible for wound healing and tissue repair across all organs and various physiological states, including cancer

Neoplasia: New, uncontrolled growth of cells that is not under physiologic control

Parenchyma: Functional tissue in plants and animals

Pathogenesis: Process by which an infection leads to disease

Prognosis: The forecast of the probable outcome or course of a disease

Proline: A cyclic, non-essential amino acid. In humans. Found in high concentrations in collagen

Proliferation: Multiplying or increasing in number. In biology, cell proliferation occurs by a process known as cell division

Proteoglycan: Proteins that are heavily glycosylated. The basic proteoglycan unit consists of a core protein with one or more covalently attached glycosaminoglycan (GAG) chain(s)

Postmenopausal: The time of life when a woman's ovaries stop producing hormones and menstrual periods stop. Occurs around age 50

Stroma: The supporting framework of an organ, a gland or other structure, usually composed of connective tissue cells, as distinguished from the parenchyma cells or tissues performing the special function of the organ or bodily part

Sebaceous: Small oil-producing gland present in the skin of mammals

Scirrhous: Scirrhous carcinomas are histologically characterized by the presence of hard, fibrous, particularly invasive tumors in which the malignant cells occur singly or in small clusters or strands in dense connective tissues

Triglyceride: A triglyceride (TG) molecule consists of a glycerol backbone esterified with three fatty acids. Triglycerides are the main constituent of vegetable and animal fats in the diet, and are the main constituent of the body's fat stores

Tubuloalveolar: Starting out as a branched tubular gland and branching further to terminate in alveoli

Versican: Large extracellular matrix proteoglycan that is present in a variety of human tissues

Viscoelastic: A material is said to be viscoelastic if it possesses both fluid and solid properties

Abbreviations

ADAMTS 1	A disintegrin and metalloproteinase thrombospondin type 1
AdipoR1	Adiponectin receptor 1
BC	Breast cancer
BI-RADS 4	Breast Imaging Reporting and Data System
BM	Basement membrane
CBCT	Cone-beam computed tomography
CD44	Cluster of differentiation 44
CS	Chondroitin sulfate
CSPG	Chondroitin sulfate proteoglycans
DCIS	Ductal carcinoma in-situ
DNA	Deoxyribonucleic acid
GAG	Glycosaminoglycan
GF	Growth factor
IAM	Independent atomic model
IDC	Invasive ductal carcinoma
ILC	Invasive lobular carcinoma
kDa	kilo-Dalton
MDA-MB-231	M.D. Anderson – Metastatic Breast 231

mRNA	Messenger ribonucleic acid
NCBI	National Center for Biotechnology Information
OBSP	Ontario Breast Screening Program
PG	Proteoglycan
RNA	Ribonucleic acid
SAXS	Small angle x-ray scatter
SVD	Singular value decomposition
TGF- β	Transforming growth factor beta
Vcan	Versican
WAXS	Wide angle x-ray scatter

Aim of the Thesis

This work aims to demonstrate a correlation between breast *cancer* biology and Wide-Angle X-ray Scatter (*WAXS*) data. The motive is to understand changes in tissue structures in the breast when cancer invades. The thesis has two major goals:

1. To characterise the breast tissue using WAXS data available in literature. This involves modeling healthy and diseased breast tissue, assuming the breast to be made up of three main components- fat, *epithelial cells* and *stroma*. This is accomplished via analysis of their WAXS signatures, namely differential linear scattering coefficient μ_s .
2. To understand structure of stroma. A stroma model, made up of *collagen*, *versican* and *elastin* is introduced. μ_s data for stroma is calculated using coherent and incoherent form factors. Dynamics of the *extracellular matrix* of the stroma in normal and diseased tissue is reviewed.

Chapter 1: Introduction

Breast cancer is the most common cancer among Canadian women (excluding non-melanoma skin cancers). Studies estimate that about 1 in 8 Canadian women develop breast cancer during their lifetime and 1 in 34 dies from it [1].

According to Canadian Cancer Society, projected estimates of breast cancer in Canada in 2022 are as follows [1]:

- 28,600 new cases of breast cancer will be diagnosed among Canadian women and 5,500 women will die from it.
- On an average, 78 Canadian women will be diagnosed with breast cancer every day and 15 will die from it (every day).

For early detection, many countries have established screening procedure that encourages women of certain risk age group to take periodical mammograms. In the province of Ontario in Canada, the Ontario Breast Screening Program (*OBSP*) recommends that most eligible women, trans and nonbinary people aged 50 to 74 should get screened with mammography every 2 years [2]. For mammography as well as modern imaging methods like breast tomosynthesis, breast CBCT etc., the primary photons carry important diagnostic information and x-ray scatter is considered as disturbance [3]. Suspicious lesions occur because of the poor contrast between a cancerous and benign lesion [3]. Tissue *biopsy* is then needed for clearer diagnosis. But sampling the wrong slice can rather lead to misdiagnosis [3]. Scientists are working on improving the process of cancer detection. One method is to use the scattering properties of different tissues to enhance the visibility of disease [4]. Some groups have demonstrated that wide angle x-ray

scatter (*WAXS*) imaging, conventionally used to study crystallographic structures ($0.5\ nm^{-1} < x < 5\ nm^{-1}$, where x is the momentum transfer variables) can be used to characterise the breast tissue [5, 6, 7, 8, 9, 10, 11, 12, 13, 14].

Other groups have stated that small angle x-ray scatter (*SAXS*) signals corresponding to collagen fibres in healthy and diseased tissue are different. Hence, SAXS may be capable of identifying areas of invasion and can provide more information at the supramolecular level ($x < 0.1\ nm^{-1}$) [15, 16, 17, 18, 19, 20, 21].

In this work, an attempt has been made to model healthy and diseased breast tissue samples by analysing WAXS signals of three of the major components of the breast: fat, epithelial cells and stroma.

While scatter studies on cells and fat have been done in the past, not much is known about stroma in the same context. Studies show that stroma promotes cancer growth and *metastasis*, making it an important component to focus on. A model for stroma majorly made up of collagen, elastin and versican is introduced. Scatter signal for stroma using Independent Atomic Model (*IAM*) approximation is compared with stroma estimation at higher values of momentum transfer.

To summarise, this work highlights the potential use of WAXS signals to characterise cancerous tissues and brings out the importance of studying stroma in *malignan*t and *benign* breast tissues.

Chapter 2: The Breast

2.1 Structure

The human breast mainly consists of the mammary glands, which are modified *tubuloalveolar apocrine* sweat glands located on the anterior thoracic wall [22]. The mammary glands consist of an epithelial bilayer embedded in adipose tissue and supported by a loose framework of dense fibrous *connective tissue*. While connective tissue gives cohesion to the ensemble of the organ, the adipose component serves as a reservoir of interstitial fluids that includes locally derived molecules for example, prolactin and those synthesized elsewhere [22].

The mammary glands are formed of 15-20 lobes, each of which has many smaller lobules, that produce milk. The milk is carried from the lobules to the nipple through thin tubes known as ducts. Each duct connects to the exterior through 0.5mm sized orifices found in the nipple [22]. Dark, pigmented, round area round the nipple is the areola which has small bumps on its surface. These are *sebaceous* glands that release oily secretions that keeps the areola and nipple lubricated.

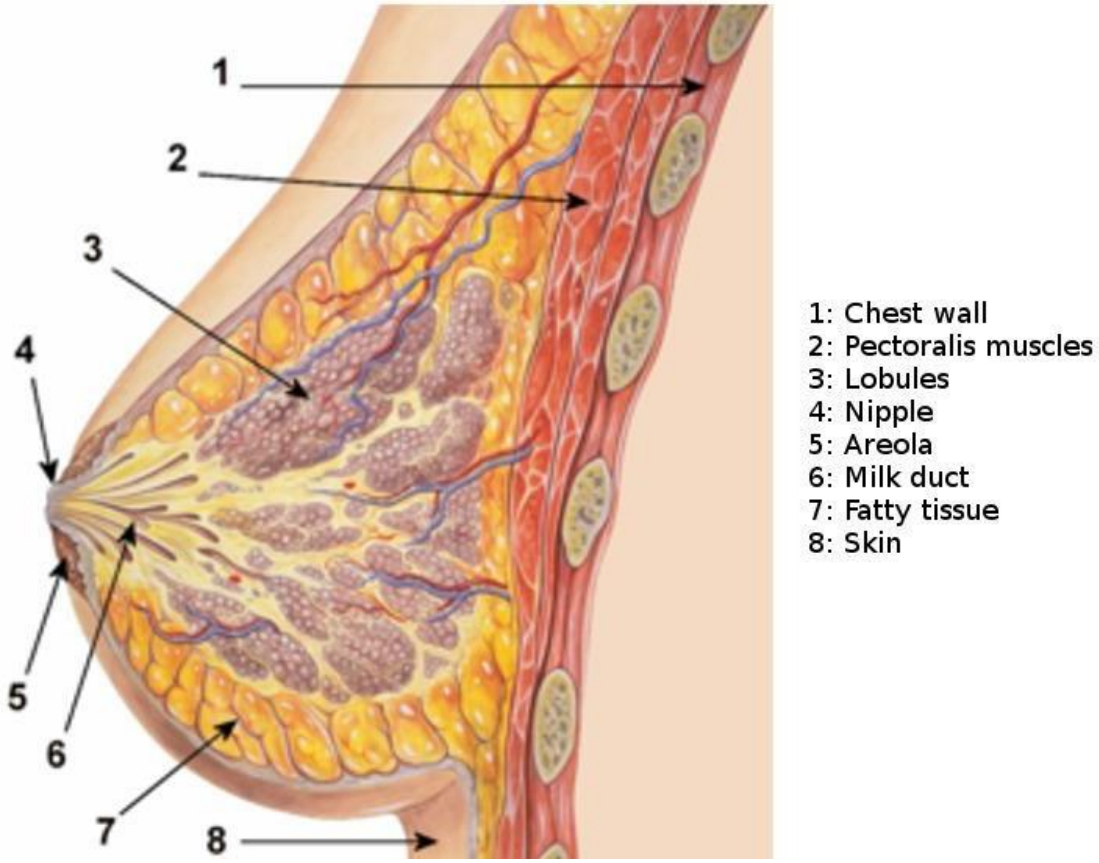


Figure 1: Anatomy of the breast *[23]*.

2.2 Building blocks of the breast: fat, glandular and fibrous tissue

Fat: Adipose or Fatty tissue fills in the spaces between glandular and fibrous tissue. It covers most of the breast from collarbone to the underarm and around the center of the ribcage, hence largely determining the breast size. Fat tissue is mainly made up of fat cells, also known as adipocytes. These cells contain fatty acids, arranged as *triglyceride* molecules [21]. The composition of fatty acids in the triglycerides vary greatly with diet and age, but they are mainly oleic and palmitic acids [21, 24]. Being a rich source of energy, it aids the development and progression of breast cancer. However, some studies suggest that the

natural secretome of breast adipose tissue is not always cancer promoting [25]. As described in the same study, when breast cancer cells are grown in conditioned media from adipose tissue explants from cancer patients, an increased expression of versican, CD44, ADAMTS1 and adipoR1 in cancer cells is seen, resulting in increased proliferation, adhesion and migration [25, 26]. While on the other hand, conditioned media from adipose tissue explants from normal breast cells decrease the migration of BC cells, implying that changes in the secretome profile of breast adipose tissue could be designated as the crosstalk between cancer cells and the adipose tissue [26].

Glandular tissue**:** Glandular tissue in general, refers to a type of epithelial tissue involved in the production and release of secretory products such as sweat, breast milk, saliva, hormones etc. Ducts and lobules in the breast are made up of glandular tissue, also known as glandular epithelium. The walls of lobules and ducts are formed by layers of epithelial cells. They are wrapped in a layer of connective tissue called *basement membrane* (mainly made up of non-fibrillar collagen) [21]. Studies show that on an average, the epithelial component of a healthy breast constitutes 10 to 15% of its overall volume [22].

Fibrous tissue**:** Also known as fibrous connective tissue, it provides support, structure and protection to the organs. Connective tissues are divided into three major groups: connective tissue proper (includes the connective tissue of stroma), specialised connective tissue and embryonic connective tissue. Found in the breast, connective tissue stroma (or simply stroma) has cellular and abundant amount of extracellular component to it. The latter, better known as the *extracellular matrix,* consists of an amorphous ground substance and

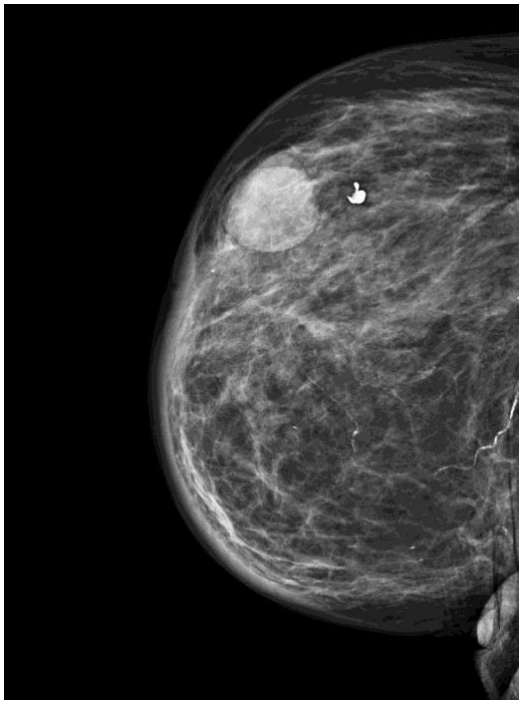
connective tissue fibers (mainly **collagen** and small amounts of **elastin**). The ground substance is a porous, well hydrated, *viscoelastic* gel made largely of *proteoglycans***.** Among the cellular components, the single most abundant type of cell is the fibroblast. Fibroblasts produce and secrete the organic components of the amorphous ground substance and fibers of the extracellular matrix [27]. Studies show stroma plays a crucial role in tumor development [28]. More about stroma is discussed in chapter 3.

2.3 Breast Disorders

Breast disorders can be benign (non-cancerous) or malignant (cancerous). Malignant tumors can be invasive and possess serious threats to the life of patient. Some of the benign and malignant conditions are discussed in this section.

2.3.1 Benign Conditions

- **Breast cysts**: Up to a quarter of breast lumps are fluid-filled *cyst*s [29]. They can be felt as a round, tender and movable lump and in most of the cases, go away without treatment. Cysts are most common in women in the age group of 30-40 years. Also, monthly hormone changes affect the size and occurrence of cysts [30]. Cysts can be of different types**: a simple cyst** which is just filled with fluid, **complicated cyst** which has something floating inside the fluid, or a **complex cystic and solid mass** which either has a solid component or thick outer layer [30]. Depending on the type a follow up exam or procedure to drain the fluid is recommended. Cysts do not increase the risk of developing breast cancer.



A complete low-density halo encircles the mass. Despite its large size it is not very dense. And the density is similar to surrounding fibroglandular tissue. Patient data: age 75 years, female. Case contributed by Dr. Henry Knipe [31].

Figure 2: Typical appearance of a benign breast cyst on a mammogram *[31]*.

- **Fibrocystic changes**: This is characterised by increase in fibrous stroma along with development of cysts. It can cause breast pain, lumpiness and tenderness.
- **Hyperplasia***:* This condition occurs from an overgrowth of epithelial cells that line mammary ducts and glands. In case of *atypical* cell growth, surgery is recommended because the benign neoplasia might develop into a malignancy [22].

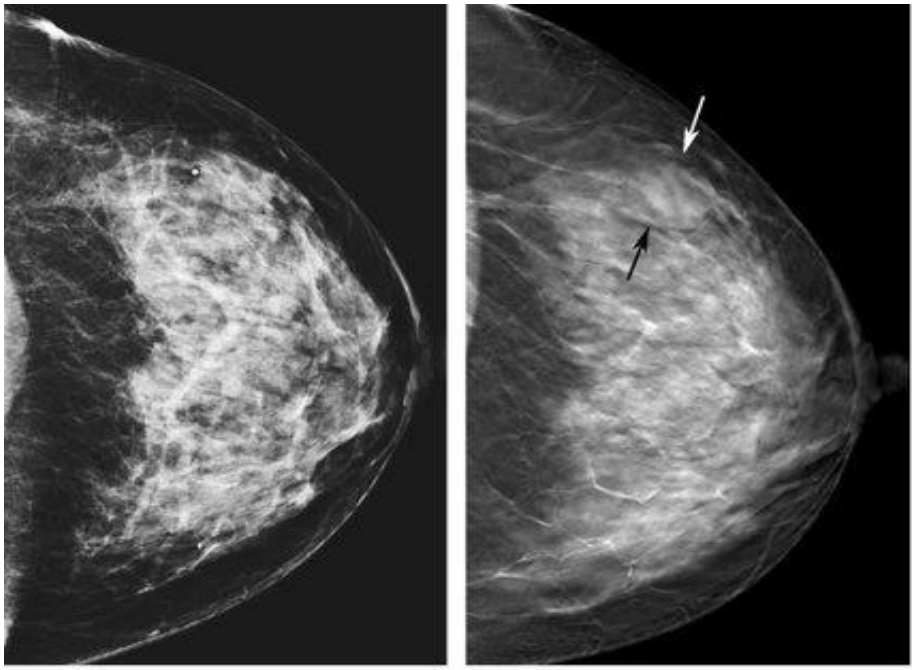


Figure 3: Fibrocystic changes and ductal hyperplasia without atypia as seen on a) mammogram and b) breast tomosynthesis image. Images taken at the same time.*[32]*.

Case of a 45-year-old woman with a palpable abnormality in the left breast. Mass is poorly depicted on the mammogram, but clearly seen on the tomosynthesis image (marked with arrows) [32].

- **Fibroadenomas**: These are smooth, rounded, solid tumors consisting of fibrous and glandular tissue that forms a mass. These lumps occur most often in women between the ages of 15 and 35 years. These are often painless and are referred to as a breast mouse due to their high mobility [33]. *Prognosis* is good as it shrinks in size over the time in most of the cases.

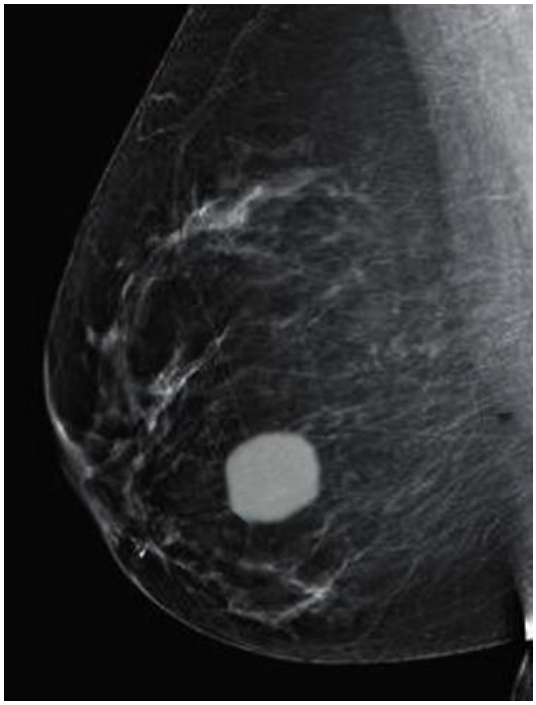
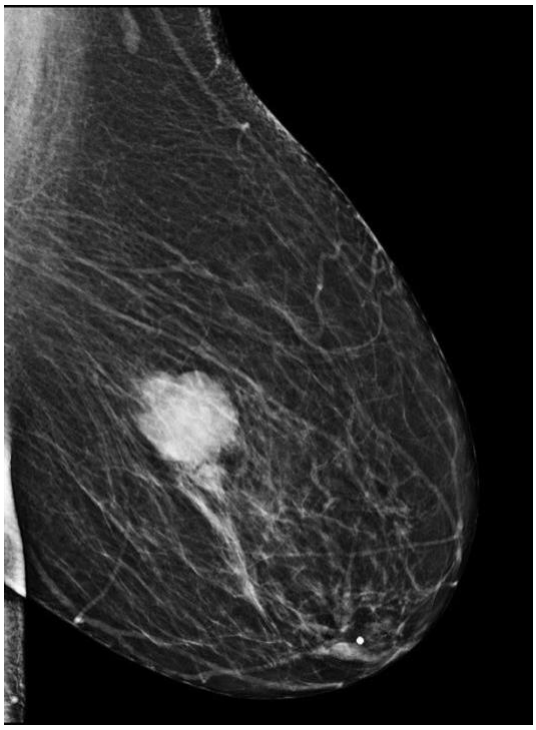


Image courtesy: S. Bhimji MD [33].

Figure 4: A typical image of fibroadenoma on a mammogram *[33]*.

- **Intraductal papilloma**: These are small wart-like growths formed inside the mammary duct near the nipple. They may cause nipple discharge and can sometimes be painful. Papilloma's presence increases the chances of developing cancer. According to a study, in 50% of the papilloma cases, the lesion looks well-defined on mammogram (BI-RADS 4) [34]. They are often recommended for ultrasound examination and biopsy.



Shown is a case of a well-defined lobulated mass of increased density in a predominantly fatty breast. Classification: BI-RADS 4. Patient data: age 70 years, female. Case contributed by Dr. Ammar Ashraf [35].

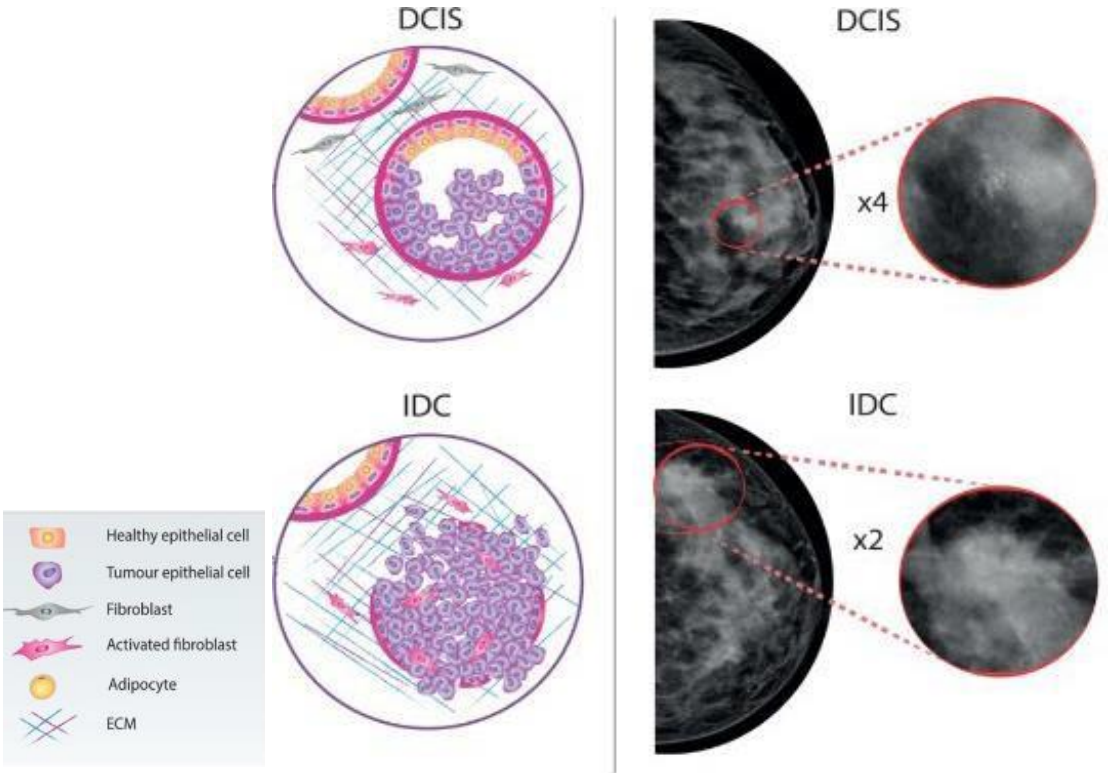
Figure 5: Intraductal papilloma of breast as it appears on mammogram *[35]*.

2.3.2 Malignant conditions

Breast cancer can begin in different parts of the breast, but most breast cancers are *carcinomas*, which are defined as tumors that start in the epithelial cells.

- **Ductal carcinoma in situ (DCIS)** is a pre-cancer that starts in a milk duct and has not grown into the rest of the breast tissue. The cancer cells are confined very locally (non-invasive) and there is no direct contact of the cancer cells with the stroma or the extracellular matrix (ECM) [22]. Survival chances are high once the tumor is removed.

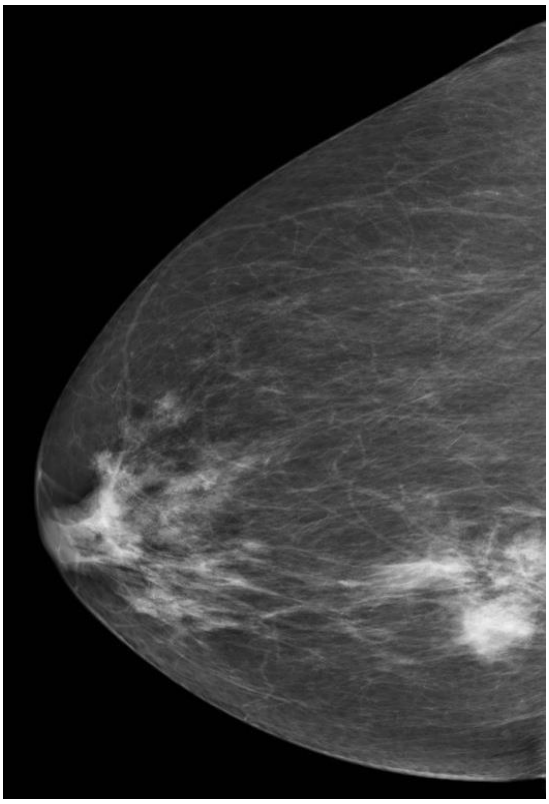
- In case of **Invasive ductal carcinoma (IDC)**, the cancer cells are in direct contact with stroma, and have already spread to the other parts of the body. IDC accounts for 70-80% of breast cancer cases.



Note that fibroblast, activated fibroblast and ECM are all part of the stroma. Adapted from Breast Mammographic Density: Stromal Implications on Breast Cancer Detection and Therapy by Fernandez-Nogueira *et. al* [18]

Figure 6: Schematic representation of mammary glands for DCIS and IDC and their representative mammograms *[18]*.

- **Invasive lobular carcinoma (ILC)** begins in the lobules and makes up about 10-15% of all breast cancers [36]. Compared with IDC, ILC is more likely to occur in older women [37]. Studies show that increase of ILC cases is associated with the use of hormone replacement therapy in *postmenopausal* patients [36]. These are challenging lesions to detect as they are very subtle and are usually found as areas of *parenchymal* asymmetry.



These lesions are relatively harder to detect, reason being that these lesions do not bring out the *scirrhous* response which is usually expected to see with IDC lesion [38].
Patient data: adult, female. Case contributed by Dr. G. Kruger [38].

Figure 7: Invasive lobular breast carcinoma on mammogram *[38]*.

Chapter 3: X-ray interactions with matter

X-rays due to their short wavelength, make an excellent tool to study structures of materials and are deemed useful in the field of crystallography, fluoroscopy and radiography. Consider a beam of x-rays with $N_0(E)$ number of photons, having energy E, incident on a material of thickness Δx. As the rays transverse through the material, some of the photons make it through without interacting and these are known as primary photons. The number of primary photons is given by [39, 40]:

$$N_p(E) = N_0(E)e^{-\mu(E).\Delta x} \quad (1)$$

where $\mu(E)$ is the linear attenuation coefficient (units cm^{-1}). Note that 'μ' describes the attenuation of the x-rays and is a function of energy. It is given by [40]:

$$\mu(E) = \mu_{pe}(E) + \mu_{pp}(E) + \mu_{coh}(E) + \mu_{inc}(E) \quad (2)$$

where μ_{pe}, μ_{pp}, μ_{coh}, μ_{inc} describes attenuation due to **photoelectric effect**, **pair production**, **coherent scattering** and **incoherent scattering** respectively. These effects are discussed below.

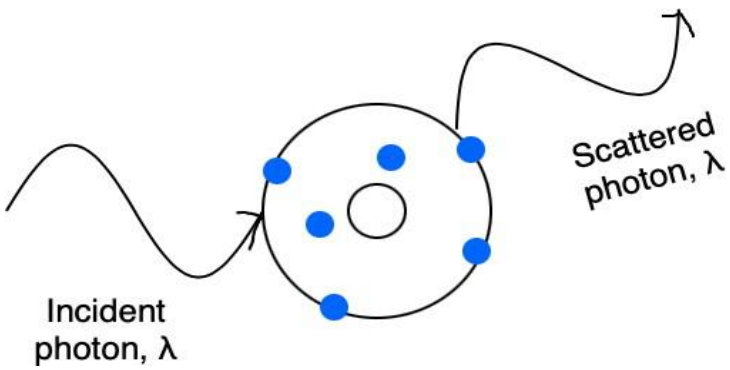
Pair production is an absorption process in which photon disappears and give rise to an electron and positron. Since a minimum photon energy of 1.022 MeV is required for pair production to occur in nuclear field, this process is not observed in diagnostic x-ray energy range [41].

When a photon hits an atom, its energy is transferred to an electron of the inner shells (absorption) and an electron known as a photoelectron is ejected from the atom. The kinetic energy of the photoelectron is given by the difference of incident photon and binding energy of the ejected electron. This phenomenon is known as the **photoelectric effect**. Note that this effect predominates over Compton effect at low photon energies [41].

Finally, the incident photon may scatter elastically or inelastically via **Coherent scattering** and **Incoherent or Compton scattering** respectively (discussed in detail in section 3.1 and 3.2). Though in mammography, scatter is considered as a nuisance as it reduces the radiographic contrast and signal-to-noise ratio, certain studies have shown that scatter signals especially the elastic scattering contains useful information about the structure of tissues and that x-ray scatter imaging can have useful application in tissue characterisation and the process of breast cancer screening [6, 8, 10, 11, 12, 13, 14, 15, 16 ,17, 19, 20, 21, 42, 43, 44, 45, 46, 47]. In the field of scatter imaging there are two regimes of interest: small-angle x-ray scatter (SAXS) for studying supramolecular structures and wide-angle x-ray scatter (WAXS) traditionally used to study crystallographic structures. This work is focused on WAXS studies and more about it is discussed in chapter 4.

3.1 Coherent Scattering

When an electromagnetic wave of wavelength λ strikes an atom, the oscillating electric field associated with the wave sets the atom in momentary vibration. The oscillating electrons emit radiation of wavelength λ, which is same as the initial radiation wavelength [39]. Hence photons retain their original energy and no energy is transferred. Acc. to Johns *et. al.* it is a cooperative phenomenon and hence is called coherent scattering [39]. The scatter signal is affected by the atomic, molecular and inter-molecular electron configuration [47].



Scattered waves from all the electrons combine and interfere to form the final wavelength which is equal to the incident photon wavelength.

Figure 8: Schematic representation of coherent scattering

As defined by Leclair in his work on x-ray scatter modeling [47], the differential coherent scatter cross section per electron, atom, or molecule, per steradian, can be written as:

$$\frac{d_K \sigma_{coh}}{\rho_e\, d\Omega} = \frac{d_e \sigma_0}{d\Omega} \mathrm{F}^2 [x, \quad (\mathbf{r})] \qquad \text{(units are } \frac{m^2}{K \cdot sr}\text{)} \qquad (3)$$

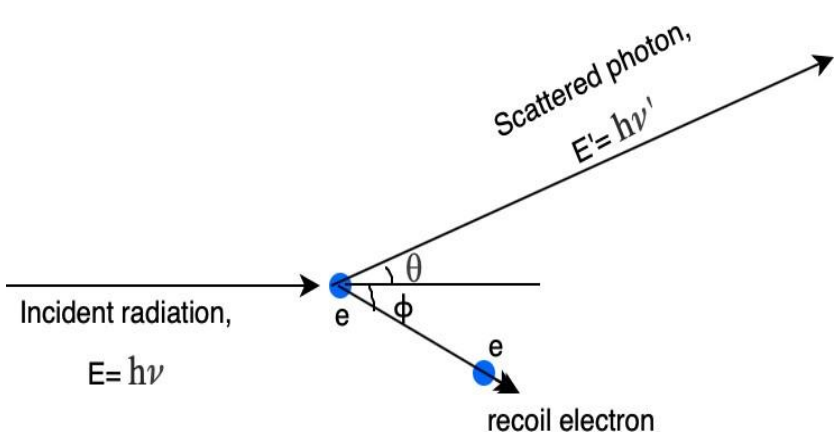
where K is the entity - electron, atom, or molecule etc. $d_e\sigma_0/d\Omega$ is the Thomson differential cross section per free electron per steradian and F is the coherent form factor that depends on the electronic distribution $\rho_e(\mathbf{r})$ and the variable x**.** Since $x = E_0 / (h.c.\sin(\theta/2))$, F is also dependent on the scattering angle θ and photon energy E_0. Note that F has no physical dimensions, and it is actually the 3-D Fourier transform of the charge distribution.

3.2 Incoherent/Compton Scattering

For this type of scatter, consider the quantum nature of the radiation and imagine the incident wave to be a stream of photons with energy $h\nu$ and momentum $h\nu/c$. When the photon hits a free electron, it scatters away taking a part of energy ($h\nu'$) at an angle θ while the rest of the energy is carried away by the recoil electron as kinetic energy E at an angle ϕ. Collision energy is conserved so the energy equation can be written as [39]:

$$h\nu = h\nu' + E \tag{4}$$

here, ν' must be less than ν, hence wavelength of the scattered waves must be longer than the wavelength of incident wave.



Adapted from Johns and Cunningham [39].

Figure 9: Schematic representation of incoherent/Compton scattering from an individual electron.

Based on the manner in which photon hits the electron, there are two main cases to discuss. When photon makes a direct hit on the electron, the electron travels exactly straight in the forward direction with $\phi = 0$, and the scattered photon will be scattered straight back with $\theta = 180°$. In this case the electron gets the maximum energy and the scattered photon its minimum energy. In another case, when the photon makes a *grazing* hit with the electron, the electron will emerge nearly at $\phi = 90°$ and scattered photon will go almost straight forward with $\theta = 0$. In this case the electron obtains almost zero energy, and the scattered photon gets the full energy of the incident photon. All intermediate angles are possible. On average, with increase in the photon energy, the fraction of energy transferred to electron as kinetic energy per collision also increases [39]. Compton scatter is the most important interaction mechanism in tissue like materials, for photons especially in the range 100 KeV to 10 MeV.

The differential cross section per *K* per steradian can be written as [47]:

$$\frac{d_k\,\sigma_{inc}}{F_{KN}\;d\Omega} = \frac{d_e\,\sigma_0}{d\Omega}\;S\,[x,Z] \qquad \left(\text{units are } \frac{m^2}{K\cdot sr}\right) \tag{5}$$

here F_{KN} is the Klein-Nishina factor that describes the probability of scattering from a free electron and S(x,Z) is the incoherent scattering function. Though S has no physical dimensions, it incorporates the effect of atomic binding on the probability of scattering [47].

Chapter 4: Wide Angle X-ray Scatter (WAXS)

4.1 Background

Wide Angle X-ray Scattering (WAXS) is a technique used for study of partially ordered materials. It is often used to characterise the crystalline structure of polymers, by measuring the interatomic spacings within the unit cell. It is sub nanometer-sized structures within the samples that cause the wide angles. Several groups have suggested WAXS as a tool to characterise breast tissue samples [4, 5, 6, 7, 8, 10, 11, 13 45, 46,]. A quantity used to characterise the WAXS properties of materials is the differential linear scattering coefficient 'μ_s' which is given by:

$$\mu_s(x) = \rho \frac{r_o^2}{2} (1+\cos^2\theta)(F_{str}^2(x) + F_{KN}(\lambda, \theta)\, S_{str}(x)) \qquad \text{(units are } sr^{-1} cm^{-1}\text{)} \qquad (6)$$

where ρ = mass density,

r_o= the classic electron radius,

F_{KN} = the Klein-Nishina function,

F_{str} (free e^-/molecular weight)$^{0.5}$ and S_{str} (free e^-/molecular weight) are the coherent form factor and incoherent scattering function of the sample,

and θ is the scatter angle.

In this thesis, WAXS data gathered by Kidane *et al*. [4] and Leclair *et al*. [3] is used to learn more about stroma and to characterise healthy, benign and malignant samples via different models.

4.2 WAXS data

In this section, some of the WAXS data present in literature which is relevant to our modeling is discussed.

The scatter data for fat [4], fibroglandular [4], stroma and cells [48] are shown in figure 10. Note that the curve for stroma was obtained via Kidane et *al.*[4] fibroglandular and fat data using the following equation:

$$\mu_s^{str}(x) = 1/0.6\ [\ \mu_s^{fib}(x) - 0.4\ \mu_s^{fat}(x)] \qquad (7)$$

where the 60% stroma and 40% contents of fibroglandular were provided by their histological study. Data for cells was taken from Leclair et *al.* [48] work on the cell model. Figure 11 compares WAXS signals from healthy, benign and malignant breast tissues as obtained by Kidane et *al*. [4].

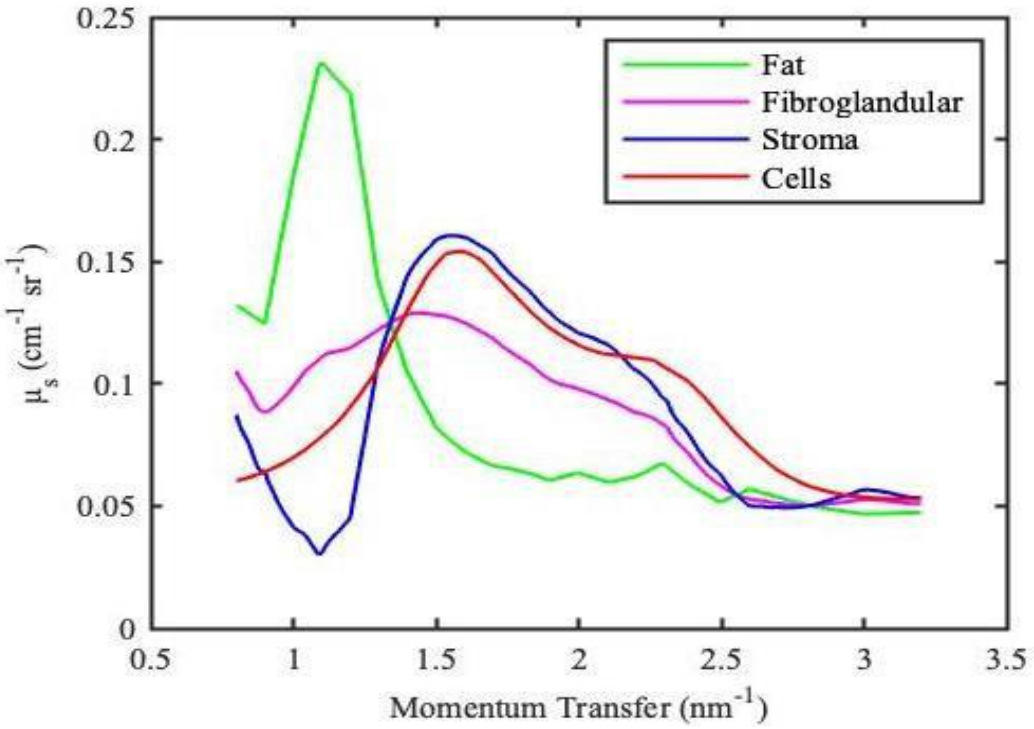


Figure 10: Scatter profiles for various tissues present in the breast *[4, 48]*.

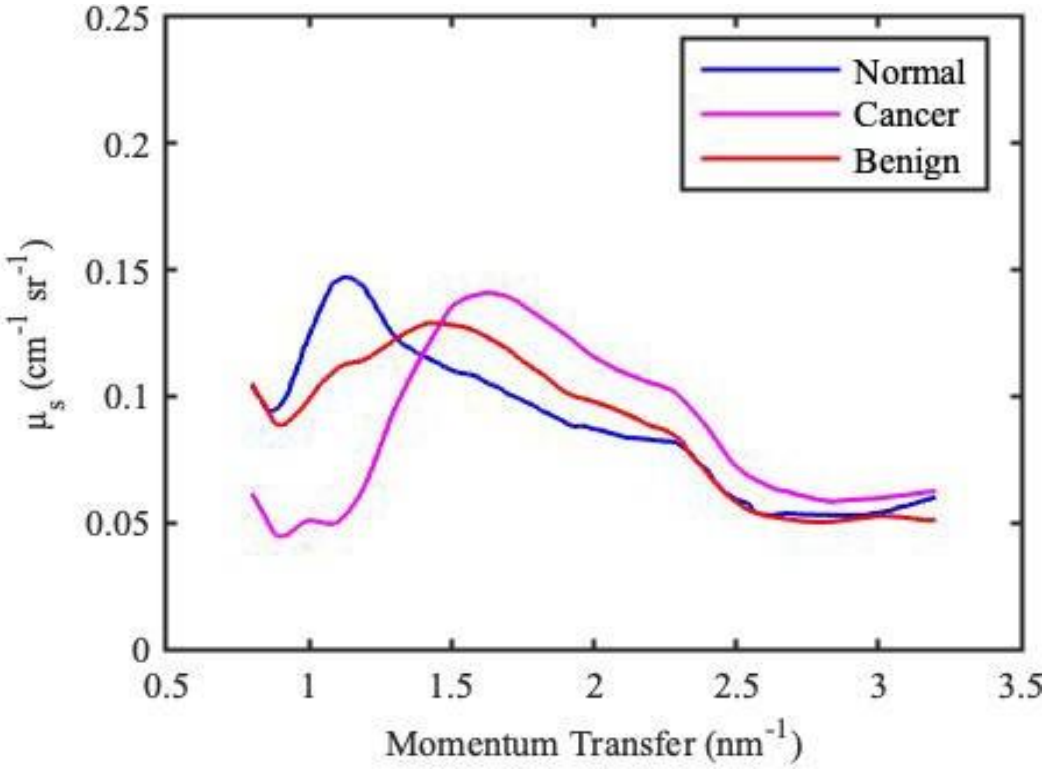


Figure 11: Scatter profile for normal, cancerous and benign breast tissue *[4]*.

The plot of μ_s for adipose tissue shows a sharper peak at lower momentum transfer values compared to other tissues. Studies have suggested that adipose tissue contains a large number of fat cells that are aligned to behave with a relatively higher degree of order, and so more evidence of diffraction effects is to be expected [49]. WAXS data for normal is similar to that of fat, indicating presence of fat in normal samples in large amount. Graphs shown above in figure 11 are fitted with the three components-fat, stroma and cells. More about the modeling and fitting is discussed in the next chapter.

Chapter 5: Data modeling and fitting

5.1 Breast tissue model

The breast tissue mainly consists of the mammary glands, embedded in adipose tissue and enfolded in connective or fibrous tissue. The mammary glands are formed by the lobules where milk is produced and stored, and by the ducts that carry milk to the nipple. A group of lobules form a lobe, and each breast has 15-20 of these lobes. Adipose tissue is situated between lobes and not within the lobules. The ductal-lobular system is lined by a specialized inner epithelial cell layer (i.e., epithelium) which has secretory and absorptive functions and is surrounded by an outer myoepithelial cell layer [50]. In this study, breast tissue is assumed to be made up of three major components-fat, fibrous tissue also known as *stroma* and epithelial cells. WAXS data of each of the component have been discussed in the previous chapter but predictions for healthy and *abnormal* tissue would be discussed in the following sections.

5.2 The cell model

In a previous study by Leclair *et al*. [48] an epithelial cell model was devised that assumed the following fractional volumes for the components in a cell: 0.771, 0.0232 and 0.2058 respectively for water, fat and other components of the cell which includes DNA, RNA, proteins and carbohydrates. The amount of carbon, oxygen, hydrogen, nitrogen, sulphur and phosphorus that exists in an epithelial cell was then calculated and the results are given in

Table 1. (refer [48] for more details on the cell model). Using this information, WAXS scatter predications from epithelial cells were derived which was shown in figure 10. Similar approach was used to get the scatter signal for stroma in the IAM regime. This would be discussed in chapter 6.

Table 1: The total number of each of the predominant atoms found in the epithelial cell, calculated by Leclair et al. [34]

Category	C	O	H	N	P	S
Water	0	1.08×10^{14}	2.16×10^{14}	0	0	0
Lipids	3.06×10^{12}	5.56×10^{11}	5.63×10^{12}	6.96×10^{10}	6.96×10^{10}	0
DNA	7.86×10^{11}	6.26×10^{11}	1.11×10^{12}	2.98×10^{11}	8.07×10^{10}	0
RNA	5.18×10^{12}	4.38×10^{12}	7.51×10^{12}	2.02×10^{12}	5.47×10^{11}	0
Proteins	1.58×10^{13}	8.76×10^{12}	3.22×10^{13}	4.56×10^{12}	0	1.43×10^{11}
Carbohydrates	2.72×10^{12}	2.72×10^{12}	5.44×10^{12}	0	0	0
Total	2.75×10^{13}	1.25×10^{14}	2.68×10^{14}	6.95×10^{12}	6.97×10^{11}	1.43×10^{11}

5.3 Fitting

As discussed, in this model the breast tissue samples are assumed to consist of varying amounts of fat, stroma (fibrous tissue) and epithelial cells. The μ_s coefficients for normal tissue and cancer/benign tumors can be expressed as a sum of the μ_s of fat, stroma and cells in the following manner:

$$\mu_s(x) = a_{fat} \cdot \mu_s^{fat}(x) + a_{str} \cdot \mu_s^{str}(x) + a_{cell} \cdot \mu_s^{cell}(x) \quad (8)$$

where $x = (1/\lambda) \sin(\theta/2)$ is the momentum transfer argument, λ is the x-ray wavelength and θ is the scatter angle. Equation (8), which looks like a three-basis function, was reduced to two-basis, making use of the argument that sum of the fractional weights of the three components a_{fat}, a_{str} and a_{cell} should be unity. Considering the set of equations within the range of x from 1 to 3.19 nm^{-1}, the fractional weight coefficients were calculated using the singular value decomposition (SVD) [51] method in MATLAB. Scatter profiles for the basis function were shown in Fig. 10. Data for normal, cancer and benign samples measured by Kidane *et al*. [4] shown in Fig. 11. were fitted with the basis function data.

Chapter 6. The Stroma Model via IAM approximation

6.1 The Stroma

Normal breast duct is made up of luminal epithelial cells covering a central *lumen* and surrounded by a layer of contractile myoepithelial cells separated from the stroma by a basement membrane (BM) [52]. The stroma comprises of cellular as well as extracellular tissue network [53]. The cellular part which includes *fibroblasts*, *myofibroblasts*, endothelial cells, adipocytes etc. is surrounded by a highly dynamic and functional *extracellular matrix* (ECM) which is mainly composed of collagens and elastin fibres embedded in a viscoelastic gel formed of proteoglycans, glycoproteins and water [53, 54]. **Collagen** is the most abundant fibrous protein with fibrillar collagen making > 80% of the ECM [55]. It constitutes the main structural element of the ECM, provides tensile strength, regulates cell adhesion, support chemotaxis and migration, and direct tissue development [56]. Regions of high breast density are associated with increased stromal collagen. Studies suggest that collagen density promotes mammary tumor initiation and progression [57]. Collagen associates with **elastin**, which is another major structural protein found in the ECM [56, 58]. Elastin fibers provide recoil to the tissues that undergo repeated stretch. Elastin-derived peptides increases the MB-231 breast cancer cell invasiveness [59]. Elastin, which has not been studied much in context of breast cancer, could therefore be an important component in the breast tumor microenvironment.

The major non fibrous component of the ECM that is the proteoglycans (PGs) are characterised by a protein portion (core protein) substituted with one or more covalently linked glycosaminoglycan (GAG) chains, resulting in a high degree of structural and functional complexity [60]. PGs are categorized on the basis of type of GAG chains attached and are of the following types: heparan sulphate PGs, chondroitin sulphate PGs (CSPGs) and dermatan sulphate PGs, or keratan sulphate PGs [54]. CSPGs, especially **versican** are the most abundant proteoglycans in the ECM of mammalian tissues [54]. Versican (Vcan) contributes to the ECM assembly and is known as a dynamic regulator of the ECM [54]. It has been reported that Vcan not only enhance breast cancer cell mammosphere and colony formation but also promote tumor formation *in vivo* [61, 62]. Increased Vcan expression is associated with cancer relapse and poor patient outcomes in various cancer types including breast cancer [61]. Studies have shown that Vcan intensifies cancer cells resistance to chemotherapy and influences markers associated with self-renewal [62]. Vcan is also increased in fibroadenoma [63]. Hence presence of Vcan cannot be neglected.

In this model, stroma is considered to be made up of 3 main components: collagen, elastin and proteoglycans specifically versican. In order to predict WAXS signature of stroma, elemental composition of each of its component is required. Therefore, structure of each of the component is discussed one by one in detail. Weight fraction percentages for collagen, versican, and elastin are assumed to be 80%, 10%, 10% respectively and the total elemental composition of stroma is then determined.

Collagen

A typical collagen molecule is formed of three **α** chains, wrapped around one another to form a triple helical structure. A single **α** chain which is about 1000 *amino acids* long, is arranged as a left-handed helix, with three amino acids per turn and with *glycine* as every third amino acid. Each **α** chain is the therefore composed of a series of Gly-X-Y repeats, where X is often *proline* and Y is often *hydroxyproline*. Out of the 28 different collagen types, collagen type I is the most abundant collagen in the human body. Collagen type I is among the breast tumor microenvironment factors, that regulates proliferation, survival, migration and invasion [64]. Type I collagen has a length of about 300nm with a radius of approximately 1.6nm. It is made up of two **α1** chains and one **α2** chain. The dataset for protein sequences of each of the chains taken from NCBI data (gene ID 1277 for COL1A1, gene ID 1278 for COL1A2) [65] was used for calculation for molecular weight of collagen and weight fractions of the constituting elements (carbon, hydrogen, oxygen, nitrogen). Note that to calculate protein molecular weight from the sequence, the molecular weights of each of the amino acids present in the sequence are added followed by subtraction of molecular weight of water for every bond created. This is because with every peptide bond formed between two amino acids, a particle of water is created and expelled. The following formula is used to calculate molecular weight from amino acid sequence:

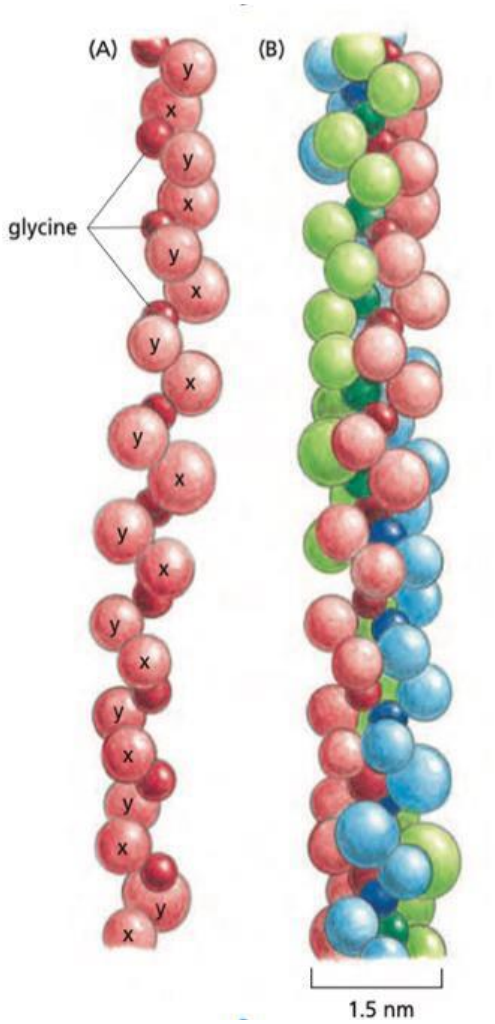
$$\text{Protein MW} = \text{Aminoacid(1) MW} + \ldots + \text{Aminoacid(n) MW} - (18.02 \cdot (n-1)) \quad (9)$$

where:

n = total number of amino acids used, and

18.02u is the molecular weight of water.

Using this formula, the molecular weight of collagen type I was calculated to be approximately 330 kDa. The weight fractions of each of the element found in the collagen molecule are tabulated in Table 2.



(A) Represents a part of a single collagen α chain, in which each amino acid is represented by a sphere. Arranged as a left-handed helix, the alpha chain is composed of a series of triplet Gly-X-Y sequences, in which X and Y can be any amino acid (although X is commonly proline and Y is commonly hydroxyproline). (B) Represents a model of part of a collagen molecule, in which three α chains, each shown in a different color, are wrapped around one another to form a triple-stranded helical structure. In collagen type I molecule, there are two α1 chains (made up of 1464 amino acids each) and one α2 chain (made up of 1366 amino acids). Glycine is the only amino acid small enough to occupy the crowded interior of the triple helix. Only a short length of the molecule is shown, the entire molecule is 300 nm long with a molecular weight of around 300 kDa.
(*Molecular biology* of *THE CELL*, sixth edition, Alberts *et al., p.* [66]*)*

Figure 12: Structure of a typical collagen molecule *[66]*.

>NP_000079.2 COL1A1 [organism=Homo sapiens] [GeneID=1277]
MFSFVDLRLLLLLAATALLTHGQEEGQVEGQDEDIPPITCVQNGLRYHDRDVWKPEPCRICVCDNGKVLCDDVICDETKNCPGAE
VPEGECCPVCPDGSESPTDQETTGVEGPKGDTGPRGPRGPAGPPGRDGIPGQPGLPGPPGPPGPPGPPGLGGNFAPQLSYGYDEKST
GGISVPGPMGPSGPRGLPGPPGAPGPQGFQGPPGEPGEPGASGPMGPRGPPGPPGKNGDDGEAGKPGRPGERGPPGPQGARGLPGT
AGLPGMKGHRGFSGLDGAKGDAGPAGPKGEPGSPGENGAPGQMGPRGLPGERGRPGAPGPAGARGNDGATGAAGPPGPTGPAG
PPGFPGAVGAKGEAGPQGPRGSEGPQGVRGEPGPPGPAGAAGPAGNPGADGQPGAKGANGAPGIAGAPGFPGARGPSGPQGPGG
PPGPKGNSGEPGAPGSKGDTGAKGEPGPVGVQGPPGPAGEEGKRGARGEPGPTGLPGPPGERGGPGSRGFPGADGVAGPKGPAGE
RGSPGPAGPKGSPGEAGRPGEAGLPGAKGLTGSPGSPGPDGKTGPPGPAGQDGRPGPPGPPGARGQAGVMGFPGPKGAAGEPGK
AGERGVPGPPGAVGPAGKDGEAGAQGPPGPAGPAGERGEQGPAGSPGFQGLPGPAGPPGEAGKPGEQGVPGDLGAPGPSGARGE
RGFPGERGVQGPPGPAGPRGANGAPGNDGAKGDAGAPGAPGSQGAPGLQGMPGERGAAGLPGPKGDRGDAGPKGADGSPGKD
GVRGLTGPIGPPGPAGAPGDKGESGPSGPAGPTGARGAPGDRGEPGPPGPAGFAGPPGADGQPGAKGEPGDAGAKGDAGPPGPA
GPAGPPGPIGNVGAPGAKGARGSAGPPGATGFPGAAGRVGPPGPSGNAGPPGPPGPAGKEGGKGPRGETGPAGRPGEVGPPGPPG
PAGEKGSPGADGPAGAPGTPGPQGIAGQRGVVGLPGQRGERGFPGLPGPSGEPGKQGPSGASGERGPPGPMGPPGLAGPPGESGR
EGAPGAEGSPGRDGSPGAKGDRGETGPAGPPGAPGAPGPVGPAGKSGDRGETGPAGPAGPVGPVGARGPAGPQGPRGDKGE
TGEQGDRGIKGHRGFSGLQGPPGPPGSPGEQGPSGASGPAGPRGPPGSAGAPGKDGLNGLPGPIGPPGPRGRTGDAGPVGPPGPPG
PPGPPGPPSAGFDFSFLPQPPQEKAHDGGRYYRADDANVVRDRDLEVDTTLKSLSQQIENIRSPEGSRKNPARTCRDLKMCHSDW
KSGEYWIDPNQGCNLDAIKVFCNMETGETCVYPTQPSVAQKNWYISKNPKDKRHVWFGESMTDGFQFEYGGQGSDPADVAIQLT
FLRLMSTEASQNITYHCKNSVAYMDQQTGNLKKALLLQGSNEIEIRAEGNSRFTYSVTVDGCTSHTGAWGKTVIEYKTTKTSRLPI
IDVAPLDVGAPDQEFGFDVGPVCFL

>NP_000080.2 COL1A2 [organism=Homo sapiens] [GeneID=1278]
MLSFVDTRTLLLLAVTLCLATCQSLQEETVRKGPAGDRGPRGERGPPGPPGRDGEDGPTGPPGPPGPPGP
PGLGGNFAAQYDGKGVGLGPGPMGLMGPRGPPGAAGAPGPQGFQGPAGEPGEPGQTGPAGARGPAGPPGKAGEDGHPGKPGRP
GERGVVGPQGARGFPGTPGLPGFKGIRGHNGLDGLKGQPGAPGVKGEPGAPGENGTPGQTGARGLPGERGRVGAPGPAGARGSD
GSVGPVGPAGPIGSAGPPGFPGAPGPKGEIGAVGNAGPAGPAGPRGEVGLPGLSGPVGPPGNPGANGLTGAKGAAGLPGVAGAPG
LPGPRGIPGPVGAAGATGARGLVGEPGPAGSKGESGNKGEPGSAGPQGPPGPSGEEGKRGPNGEAGSAGPPGPPGLRGSPGSRGLP
GADGRAGVMGPPGSRGASGPAGVRGPNGDAGRPGEPGLMGPRGLPGSPGNIGPAGKEGPVGLPGIDGRPGPIGPAGARGEPGNIG
FPGPKGPTGDPGKNGDKGHAGLAGARGAPGPDGNNGAQGPPGPQGVQGGKGEQGPPGPPGFQGLPGPSGPAGEVGKPGERGLH
GEFGLPGPAGPRGERGPPGESGAAGPTGPIGSRGPSGPPGPDGNKGEPGVVGAVGTAGPSGPSGLPGERGAAGIPGGKGEKGEPGL
RGEIGNPGRDGARGAPGAVGAPGPAGATGDRGEAGAAGPAGPAGPRGSPGERGEVGPAGPNGFAGPAGAAGQPGAKGERGAKG
PKGENGVVGPTGPVGAAGPAGPNGPPGPAGSRGDGGPPGMTGFPGAAGRTGPPGPSGISGPPGPPGPAGKEGLRGPRGDQGPVGR
TGEVGAVGPPGFAGEKGPSGEAGTAGPPGTPGPQGLLGAPGILGLPGSRGERGLPGVAGAVGEPGPLGIAGPPGARGPPGAVGSPG
VNGAPGEAGRDGNPGNDGPPGRDGQPGHKGERGYPGNIGPVGAAGAPGPHGPVGPAGKHGNRGETGPSGPVGPAGAVGPRGPS
GPQGIRGDKGEPGEKGPRGLPGLKGHNGLQGLPGIAGHHGDQGAPGSVGPAGPRGPAGPSGPAGKDGRTGHPGTVGPAGIRGPQ
GHQGPAGPPGPPGPPGPPGVSGGGYDFGYDGDFYRADQPRSAPSLRPKDYEVDATLKSLNNQIETLLTPEGSRKNPARTCRDLRLS
HPEWSSGYYWIDPNQGCTMDAIKVYCDFSTGETCIRAQPENIPAKNWYRSSKDKKHVWLGETINAGSQFEYNVEGVTSKEMATQ
LAFMRLLANYASQNITYHCKNSIAYMDEETGNLKKAVILQGSNDVELVAEGNSRFTYTVLVDGCSKKTNEWGKTIIEYKTNKPSR
LPFLDIAPLDIGGADQEFFVDIGPVCFK

Collagen **α1** and **α2** chain protein sequence [65]. Proteins are built up from 20 amino acids. Each alphabet in the sequence corresponds to a specific amino acid. There are 4294 residues in Collagen I sequence.

Elastin

Elastin protein is made up of cross-linked soluble tropoelastin molecules, that are secreted from smooth muscle cells and fibroblasts. These tropoelastin molecules are catalyzed by lysyl oxidase to make larger globules that assemble onto a proteinaceous scaffold made up of various proteins including fibrillins, fibulins and glycoproteins. This scaffold acts as a base onto which elastin protein aggregates and assembles [58]. The focus of study is the single elastin molecule without the scaffolds. This protein is rich in hydrophobic amino acids such as glycine and proline, which form mobile hydrophobic regions bounded by crosslinks between lysine residues. Protein sequence data shown below was taken from NCBI data (gene ID: 2006) and a similar set of calculations as for collagen were done. Using equation (9) the molecular weight of the single elastin molecule was calculated to be ~50kDa and the weight fractions of each of the element found in elastin are tabulated in Table 2.

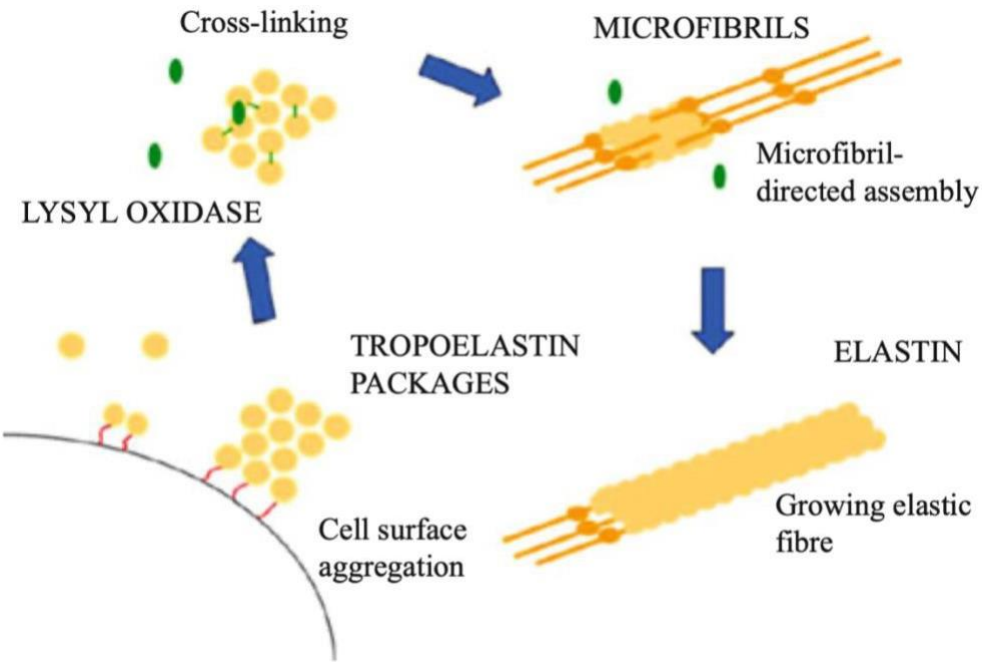


Figure 13: Assembly mechanism of elastic fibers from tropoelastin molecules *[58]*.

>NP_000492.2 ELN [organism=Homo sapiens] [GeneID=2006] [isoform=a precursor]
MAGLTAAAPRPGVLLLLLSILHPSRPGGVPGAIPGGVPGGVFYPGAGLGALGGGALGPGGKPLKPVPGGLAGAGLGAGLGAFPAV
TFPGALVPGGVADAAAAYKAAKAGAGLGGVPGVGGLGVSAGAVVPQPGAGVKPGKVPGVGLPGVYPGGVLPGARFPGVGVLP
GVPTGAGVKPKAPGVGGAFAGIPGVGPFGGPQPGVPLGYPIKAPKLPGGYGLPYTTGKLPYGYGPGGVAGAAGKAGYPTGTGVG
PQAAAAAAAAKAAAKFGAGAAGVLPGVGGAGVPGVPGAIPGIGGIAGVGTPAAAAAAAAAAKAAKYGAAAGLVPGGPGFGPGV
VGVPGAGVPGVGVPGAGIPVVPGAGIPGAAVPGVVSPEAAAKAAAKAAKYGARPGVGVGGIPTYGVGAGGFPGFGVGVGGIPG
VAGVPGVGGVPGVGGVPGVGISPEAQAAAAAKAAKYGVGTPAAAAAKAAAKAAQFGLVPGVGVAPGVGVAPGVGVAPGVGL
APGVGVAPGVGVAPGVGVAPGIGPGGVAAAAKSAAKVAAKAQLRAAAGLGAGIPGLGVGVGVPGLGVGAGVPGLGVGAGVPG
FGAVPGALAAAKAAKYGAAVPGVLGGLGALGGVGIPGGVVGAGPAAAAAAAKAAAKAAQFGLVGAAGLGGLGVGGLGVPGV
GGLGGIPPAAAAKAAKYGAAGLGGVLGGAGQFPLGGVAARPGFGLSPIFPGGACLGKACGRKRK

Elastin protein sequence [65].

Versican

Vcan is a large chondroitin sulphate PG which is secreted and incorporated into the ECM. It is transcribed from a single gene which is localised on chromosome 5q 12-14 in human genome and extends over 90 kb which is divided into 15 exons ranging in size from 76 to 5262 bp [54, 67, 68]. The alternative mRNA-splicing of these exons gives the four classical isoforms: V0, V1, V2 and V3 distinguished by different core middle regions [54, 69]. Versican is structurally composed of an N-terminal G1 domain, a CS attachment region, and a C terminus (or G3) selectin like domain. The G3 domain is composed of 2 EGF-like repeats, a lectin-like motif, and a complement binding protein-like motif. V1 isoform has shown increased expression in various kinds of tumors. Hence V1 isoform is used in modeling. Using protein sequence data from NCBI Gene ID: 1462, similar calculations as done for collagen and versican were performed. The molecular weight of the protein core was estimated to be 310 kDa and the elemental composition is tabulated in table 2.

>NP_004376.2 VCAN [organism=Homo sapiens] [GeneID=1462] [isoform=1 precursor]
MFINIKSILWMCSTLIVTHALHKVKVGKSPPVRGSLSGKVSLPCHFSTMPTLPPSYNTSEFLRIKWSKIEVDKNGKDLKETTVLVAQN
GNIKIGQDYKGRVSVPTHPEAVGDASLTVVKLLASDAGLYRCDVMYGIEDTQDTVSLTVDGVVFHYRAATSRYTLNFEAAQKACL
DVGAVIATPEQLFAAYEDGFEQCDAGWLADQTVRYPIRAPRVGCYGDKMGKAGVRTYGFRSPQETYDVYCYVDHLDGDVFHLTV
PSKFTFEEAAKECENQDARLATVGELQAAWRNGFDQCDYGWLSDASVRHPVTVARAQCGGGLLGVRTLYRFENQTGFPPPDSRFD
AYCFKPKEATTIDLSILAETASPSLSKEPQMVSDRTTPIIPLVDELPVIPTEFPPVGNIVSFEQKATVQPQAITDSLATKLPTPTGSTKKP
WDMDDYSPSASGPLGKLDISEIKEEVLQSTTGVSHYATDSWDGVVEDKQTQESVTQIEQIEVGPLVTSMEILKHIPSKEFPVTETPLV
TARMILESKTEKKMVSTVSELVTTGHYGFTLGEEDDEDRTLTVGSDESTLIFDQIPEVITVSKTSEDTIHTHLEDLESVSASTTVSPLIM
PDNNGSSMDDWEERQTSGRITEEFLGKYLSTTPFPSQHRTEIELFPYSGDKILVEGISTVIYPSLQTEMTHRRERTETLIPEMRTDTYTD
EIQEEITKSPFMGKTEEEVFSGMKLSTSLSEPIHVTESSVEMTKSFDFPTLITKLSAEPTEVRDMEEDFTATPGTTKYDENITTVLLAHG
TLSVEAATVSKWSWDEDNTTSKPLESTEPSASSKLPPALLTTVGMNGKDKDIPSFTEDGADEFTLIPDSTQKQLEEVTDEDIAAHGKF
TIRFQPTTSTGIAEKSTLRDSTTEEKVPPITSTEGQVYATMEGSALGEVEDVDLSKPVSTVPQFAHTSEVEGLAFVSYSSTQEPTTYVD
SSHTIPLSVIPKTDWGVLVPSVPSEDEVLGEPSQDILVIDQTRLEATISPETMRTTKITEGTTQEEFPWKEQTAEKPVPALSSTAWTPKE
AVTPLDEQEGDGSAYTVSEDELLTGSERVPVLETTPVGKIDHSVSYPPGAVTEHKVKTDEVVTLTPRIGPKVSLSPGPEQKYETEGSS
TTGFTSSLSPFSTHITQLMEETTTEKTSLEDIDLGSGLFEKPKATELIEFSTIKVTVPSDITTAFSSVDRLHTTSAFKPSSAITKKPPLIDRE
PGEETTSDMVIIGESTSHVPPTTLEDIVAKETETDIDREYFTTSSPPATQPTRPPTVEDKEAFGPQALSTPQPPASTKFHPDINVYIIEVRE
NKTGRMSDLSVIGHPIDSESKEDEPCSEETDPVHDLMAEILPEFPDIIEIDLYHSEENEEEEECANATDVTTTPSVQYINGKHLVTTVP
KDPEAAEARRGQFESVAPSQNFSDSSESDTHPFVIAKTELSTAVQPNESTETTESLEVTWKPETYPETSEHFSGGEPDVFPTVPFHEEF
ESGTAKKGAESVTERDTEVGHQAHEHTEPVSLFPEESSGEIAIDQESQKIAFARATEVTFGEEVEKSTSVTYTPTIVPSSASAYVSEEEA
VTLIGNPWPDDLLSTKESWVEATPRQVVELSGSSSIPITEGSGEAEEDEDTMFTMVTDLSQRNTTDTLITLDTSRIITESFFEVPATTIYP
VSEQPSAKVVPTKFVSETDTSEWISSTTVEEKKRKEEEGTTGTASTFEVYSSTQRSDQLILPFELESPNVATSSDSGTRKSFMSLTTPTQ
SEREMTDSTPVFTETNTLENLGAQTTEHSSIHQPGVQEGLTTLPRSPASVFMEQGSGEAAADPETTTVSSFSLNVEYAIQAEKEVAGT
LSPHVETTFSTEPTGLVLSTVMDRVVAENITQTSREIVISERLGEPNYGAEIRGFSTGFPLEEDFSGDFREYSTVSHPIAKEETVMMEGS
GDAAFRDTQTSPSTVPTSVHISHISDSEGPSSTMVSTSAFPWEEFTSSAEGSGEQLVTVSSSVVPVLPSAVQKFSGTASSIIDEGLGEVG
TVNEIDRRSTILPTAEVEGTKAPVEKEEVKVSGTVSTNFPQTIEPAKLWSRQEVNPVRQEIESETTSEEQIQEEKSFESPQNSPATEQTIF
DSQTFTETELKTTDYSVLTTKKTYSDDKEMKEEDTSLVNMSTPDPDANGLESYTTLPEATEKSHFFLATALVTESIPAEHVVTDSPIK
KEESTKHFPKGMRPTIQESDTELLFSGLGSGEEVLPTLPTESVNFTEVEQINNTLYPHTSQVESTSSDKIEDFNRMENVAKEVGPLVSQ
TDIFEGSGSVTSTTLIEILSDTGAEGPTVAPLPFSTDIGHPQNQTVRWAEEIQTSRPQTITEQDSNKNSSTAEINETTTSSTDFLARAYGF
EMAKEFVTSAPKPSDLYYEPSGEGSGEVDIVDSFHTSATTQATRQESSTTFVSDGSLEKHPEVPSAKAVTADGFPTVSVMLPLHSEQN
KSSPDPTSTLSNTVSYERSTDGSFQDRFREFEDSTLKPNRKKPTENIIIDLDKEDKDLILTITESTILEILPELTSDKNTIIDIDHTKPVYEDI
LGMQTDIDTEVPSEPHDSNDESNDDSTQVQEIYEAAVNLSLTEETFEGSADVLASYTQATHDESMTYEDRSQLDHMGFHFTTGIPAP
STETELDVLLPTATSLPIPRKSATVIPEIEGIKAEAKALDDMFESSTLSDGQAIADQSEIIPTLGQFERTQEEYEDKKHAGPSFQPEFSSG
AEEALVDHTPYLSIATTHLMDQSVTEVPDVMEGSNPPYYTDTTLAVSTFAKLSSQTPSSPLTIYSGSEASGHTEIPQPSALPGIDVGSS
VMSPQDSFKEIHVNIEATFKPSSEEYLHITEPPSLSPDTKLEPSEDDGKPELLEEMEASPTELIAVEGTEILQDFQNKTDGQVSGEAIKM
FPTIKTPEAGTVITTADEIELEGATQWPHSTSASATYGVEAGVVPWLSPQTSERPTLSSSPEINPETQAALIRGQDSTIAASEQQVAARI
LDSNDQATVNPVEFNTEVATPPFSLLETSNETDFLIGINEESVEGTAIYLPGPDRCKMNPCLNGGTCYPTETSYVCTCVPGYSGDQCE
LDFDECHSNPCRNGATCVDGFNTFRCLCLPSYVGALCEQDTETCDYGWHKFQGQCYKYFAHRRTWDAAERECRLQGAHLTSILSH
EEQMFVNRVGHDYQWIGLNDKMFEHDFRWTDGSTLQYENWRPNQPDSFFSAGEDCVVIIWHENGQWNDVPCNYHLTYTCKKGT
VACGQPPVVENAKTFGKMKPRYEINSLIRYHCKDGFIQRHLPTIRCLGNGRWAIPKITCMNPSAYQRTYSMKYFKNSSSAKDNSINTS
KHDHRWSRRWQESRR

Protein sequence for versican [65].

Molecular masses of tissues are usually unknown but considering a biopsy sized sample of stromal tissue (d=2mm, t=5mm) with density as 1.027 g/cm^3, weight fraction of each element can be determined. To do this, number of collagen, versican and elastin molecules in the given sample are first calculated using the equation:

$$N_{col} = (N_A \cdot M_{col})/MM_{col} \quad (10)$$

Where N_A is Avogadro's number, M_{col} is weight of collagen in the given stroma sample ($M_{col} = 0.8M_{str}$) and MM_{col} is the molecular mass of collagen. N_{ela} and N_{ver} can be calculated in a

similar way. Now that number of molecules of the three stromal components are known, the total number of C, H, O, N, S atoms present in the sample can be given by:

$$N_C = (N_{col} \cdot 17551) + (N_{ver} \cdot 16297) + (N_{ela} \cdot 2868) \quad (11)$$

where 17551, 16297, 2868 are the number of carbon atoms in one molecule of collagen, versican and elastin fibril respectively. Number of atoms of the other four elements can be calculated in a similar fashion. The final step would be changing this into elemental weight fractions which is simply given by:

$$wC = (N_C \cdot A_C) / (N_C \cdot A_C + N_H \cdot A_H + N_N \cdot A_N + N_O \cdot A_O + N_S \cdot A_S) \quad (12)$$

where A_i is the atomic mass of the element. The final results are tabulated in the last row of Table 2.

Table 2: Components of stroma and their elemental composition.

Component	**Weight fraction (w)**				
	C	H	N	O	S
Collagen	0.639	0.057	0.230	0.066	0.008
Elastin	0.684	0.078	0.221	0.015	0.002
Versican	0.628	0.061	0.190	0.112	0.009
Stroma	0.643	0.060	0.225	0.065	0.007

Studies show that on an average, about 60-70% of the fibroglandular tissue content is water [70]. So, it is safe to say that fibroglandular tissue is made up of approximately 65% epithelial cells and 35% stromal composition. To check the validity of our stroma model,

calculation for elemental composition of fibroglandular is done using the above stated ratio of the two components. This is then compared with Poletti *et al.* glandular tissue composition [46]. The comparison is shown in table 3. Note that the two values are close yet there is some difference as it is unclear if by glandular tissue samples taken by Poletti included the fibrous part of the breast.

Table 3: Comparison of elemental composition of Poletti et al. glandular tissue and Fibroglandular tissue calculated in this study.

Element	**Weight fraction**	
	Poletti's Glandular tissue	Fibroglandular tissue (65%cells+35% stroma)
C	0.184	0.225
H	0.093	0.094
N	0.044	0.079
O	0.679	0.599
S	-	0.003

6.2 Independent Atomic Model (IAM)

At higher values of x, the coherent form factor F for stroma can be approximated by a free atom form factor F_{free}. According to IAM approximation, scatter effect from each atom is independent and effects of interference between electrons of different atoms in the molecule can be neglected. This holds true only at higher x values. For example: the square of free form factor for water can be written as [47]

$$F^2_{free,\,w} = 2F^2_H + F^2_O \quad \text{(free electron/molecule)} \quad (13)$$

where F_H and F_O are atomic form factors for hydrogen and oxygen calculated by Hubbell et. *al.* [71]

Similarly, for collagen:

$$F^2_{free,\,col} = 17551 \cdot F^2_C + 18753 \cdot F^2_H + 5411 \cdot F^2_N + 1358 \cdot F^2_O + 81 \cdot F^2_S \quad \text{(free electron/molecule)} \qquad (14)$$

Hence, for stroma:

$$F^2_{free,\,str} = N_{col} \cdot F^2_{col,\,free} + N_{ver} \cdot F^2_{ver,\,free} + N_{ela} \cdot F^2_{ela,free} \quad \text{(free electron/stroma matrix)} \qquad (15)$$

$S_{free,\,str}$ can be written as:

$$S_{free,\,str} = N_{col} \cdot S_{col,free} + N_{ver} \cdot S_{ver,free} + N_{ela} \cdot S_{ela,free} \quad \text{(free electron/stroma matrix)} \qquad (16)$$

Graphs for F^2, F^2/M and S for each of the component is shown below. Note that comparing F^2/M is more useful due to major differences in the molecular masses of the the three components especially elastin. It is clear that once normalised, the coherent signals from collagen and versican are almost the same, however for elastin one can see a small difference. Using the coherent and incoherent form factors calculated for stroma, one can obtain its scatter signal μ_s in the IAM regime using the following relation:

$$\mu_s^{str}(x) = \frac{\rho_m}{2} \left(F^2_{free,\,str}(x) + F_{KN}(\lambda, \theta) \cdot S_{free,str}(x) \right) \qquad (17)$$

- where ρ_m is no. of molecules per unit volume, and can be calculated using the formula $\rho_m = (\rho \cdot N_A)/\text{Molar Mass}$

  r_o= 2.8179e-13 cm, which is the classic electron density

$\theta = 6°$ (Kidane conducted the experiment at this angle)

F_{KN} is the Klein Nishina function

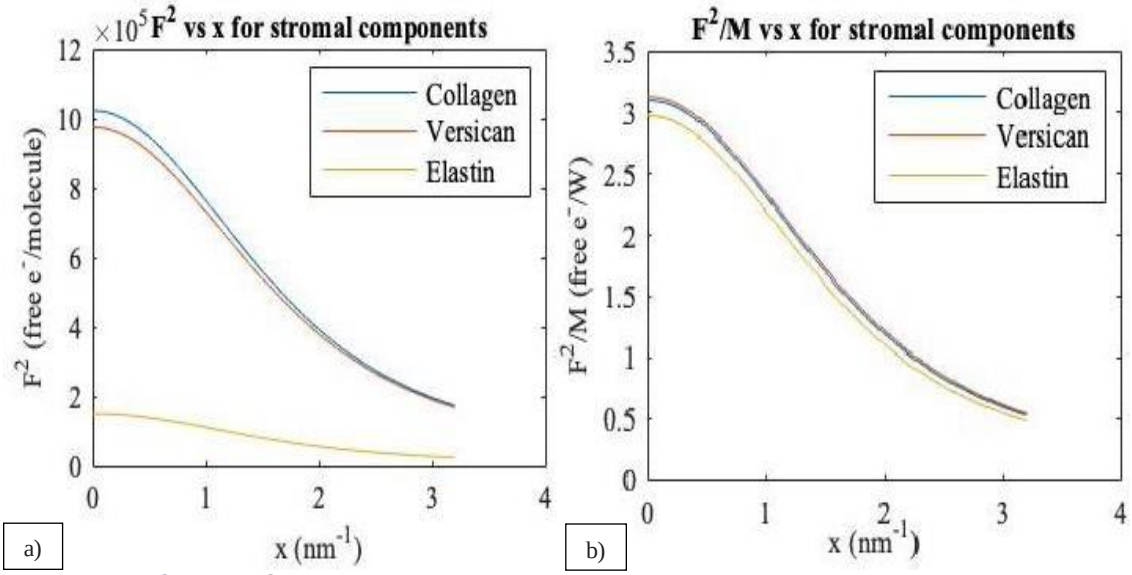


Figure 14: a) F^2 and b) F^2/M graphs for collagen, versican, elastin.

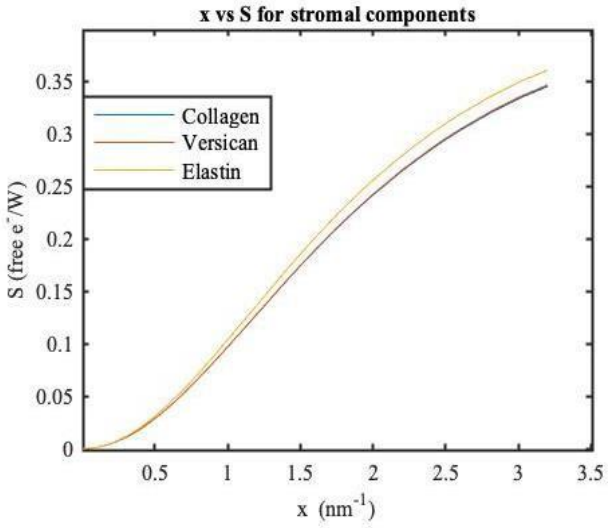


Figure 15: S for the three stromal components.

Chapter 7: Results and review

Results section has been divided into two parts. Section 7.1 discusses the fitting results from breast tissue modeling and its co-relation with cancer biology. The results for stroma modelling via IAM is discussed in section 7.2. Finally, to highlight the importance of ECM, a review of ECM dynamics in normal and diseased breast tissue is presented.

7.1 Results: Breast tissue modeling

Table 4: Fitting results from breast tissue model.

Tissue Type	**Percentage by volume (%)**		
	Fat	Stroma	Cells
Normal	56.5	35.4	8.19
Benign	25.6	67.3	7.11
Malignant	3.44	52.9	43.7

Uncontrolled proliferation of cells is a hallmark of cancer, and our results suggest the same. Percentage of cells in normal and benign is approximately the same, ranging from 7-8% of the total composition, but increases dramatically to 43.7% in the cancerous sample. Another interesting finding is that the fat percentage decreases substantially from 56.5% in the normal tissue to 3.44% in the malignant sample while the stromal percentage increases. The stroma of invasive carcinoma, i.e., the connective tissue that supports the malignant cells, varies, being highly fibroblastic to densely hyaline and contains collagen, extracellular mucin and elastic tissue in varying amounts [8]. This tissue replaces the fat as the tumor invades it [8]. So, carcinoma is left with the lack of fat within its mass and hence gives rise to low

intensity scattered signals [14]. Benign diseases like fibroadenoma are characterized by a surge in 'fibrous' tissue, hence we expect stroma to show an increase in benign diseases. Table 4 states that 67.3% of the total volume of a benign breast sample is merely connective tissue.

7.2 Results: Scatter signal for stroma via IAM

Using equation (17), the differential linear scattering coefficient 'μ_s' for stroma was determined and is shown in figure 16. The scatter data for stroma via IAM (composition: 80% collagen, 10% elastin and 10% versican) agrees with stroma data estimated using Kidane's data for fibroglandular and fat, mainly at the higher values of momentum transfer values. This is because IAM fails at the lower regime due to interference phenomenon.

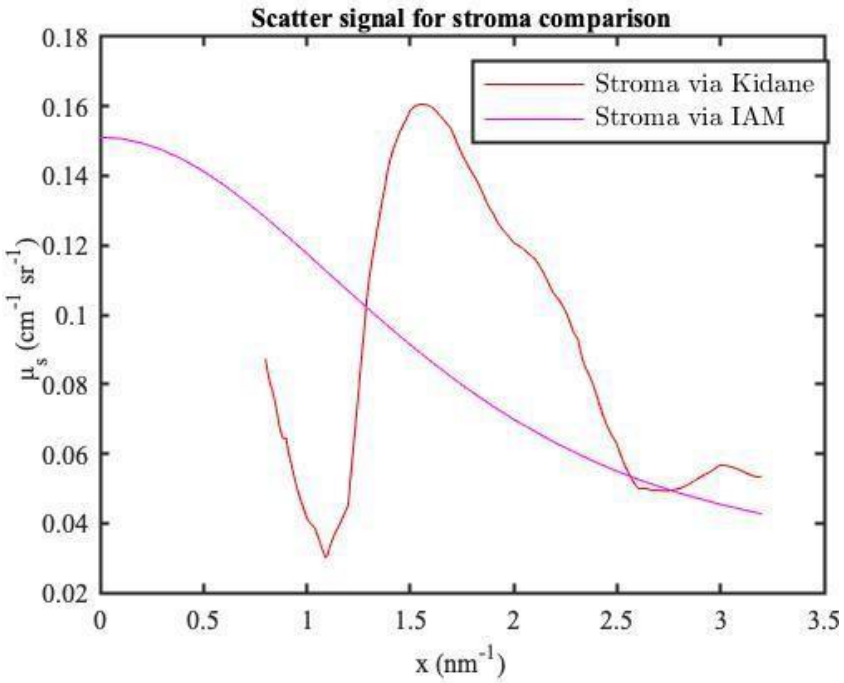


Figure 16: Comparison between μ_s for stroma obtained via IAM and estimation using Kidane's data

Scatter signals for stroma using several combinations of constituting percentages of the three components were computed in order to test the sensitivity of the model. Corresponding μ_s values were then compared with μ_s of the original model (collagen 80%, elastin and versican 10% each). Results are summarised in table 5.

Table 5: shows how values of μs change (minimum to maximum) when composition of stroma is varied.

Constituting percentage (%)			**Percentage change (%)**
Collagen	Elastin	Versican	
50	25	25	0.24 to 0.50
33.34	33.33	33.33	0.37 to 0.77
10	10	80	-0.65 to -1.42
10	80	10	2.41 to 3.49

It is observed that the scatter values obtained using IAM are not very sensitive to the change in composition of stroma. In most of the cases, the percentage change with reference to the original model is less than 1%. In the event in which elastin is considered to be the major constituent, a difference of 2.41% to 3.49% is seen. To conclude, the newly defined stroma model does not seem to be very sensitive to the percentage composition of the three building blocks at the IAM level. However, there might be a different picture if the scatter signals are obtained experimentally.

7.3 Review: ECM dynamics

ECM in normal scenario: As it can be seen from the fitting results, normal tissue is about a 50-50 mixture of fat, and stroma and cells together. The normal glandular epithelial tissue is made from of a layer of epithelial cells having apical-basal polarity, where the basal side

contacts the basement membrane, and the apical side is facing opposite to the fluid filled lumen [57]. In a compliant tissue, ECM is made up of a relaxed meshwork of fibers like collagens and elastin surrounded by and embedded in a hydrogel of glycosaminoglycan-chain-containing PGs [57]. This network allows the ECM to resist different types of tensile stresses. Binding of hydrated glycosaminoglycan network to the fibrous ECM molecules helps a healthy functional tissue to resist compressive stresses [57].

Fibrosis: a benign condition: During *pathogenesis*, the composition of the ECM is largely altered [53]. Studies show that altered ECM is not only a consequence but also a driver of fibrosis [72]. Fibrosis progression involves both cell-intrinsic and ECM-driven mechanisms. Excessive ECM production and abnormal ECM turnover are two vital signs of progressive fibrosis [72]. Increase in tissue's elastic modulus i.e., stiffness is another hallmark of fibrotic tissue. Fibrosis has been likened to atypical wound healing in which there is progression rather than resolution of scarring [73]. Early phase of wound healing involves formation of a provisional ECM. Fibroblasts occupy this matrix and proliferate in response to activators elaborators by leukocytes that have migrated into the wound and are retained by the ECM [73]. Along with this appears myofibroblasts, which are specialized form of fibroblast whose differentiation is majorly driven by cytokines, such as transforming growth factor-β and by mechanical tension. In the next phase of scar tissue shrinkage, myofibroblasts undergo apoptosis, forming a collagen-rich, cell-deficit scar [73]. This supports our fitting results for benign shown in table 4.

Tumor: Cancer is the loss of tissue organisation and abnormal behaviour of cellular components. The ECM of the breast becomes progressively stiffer and more collagen rich during tumor progression in a process called '*desmoplasia*' [52]. The stiffening of tumors is prompted by ECM deposition and remodeling by resident fibroblasts, and by increased contractility of the transformed epithelium [56]. Moreover, chemokines and GF's induce inflammation. This inflammation potentiates stromal fibroblast activation and induces their transdifferentiation into myofibroblasts, thus promoting tissue desmoplasia. Mammary tumor progression is usually classified into two major stages (1) ductal carcinoma *in situ* (DCIS), that is the premalignant stage and (2) invasive ductal carcinoma (IDC), characterised by an invasive phenotype, which includes breaching of the BM and migration of the tumor cells into neighbouring tissues. Studies show that as breast tumors progress from DCIS to IDC, the ECM undergoes increased collagen fiber linearization and thickening due to deposition and cross linking of collagen [52]. While in DCIS, linearized and thickened collagen I fibers are oriented adjacent to tumor boundary, it is aligned perpendicular to the tumor boundary in IDC, forming migration tracks for invasive tumor cells to exit the tumor tissue and enter the blood stream [52]. Increased presence of versican is also observed.

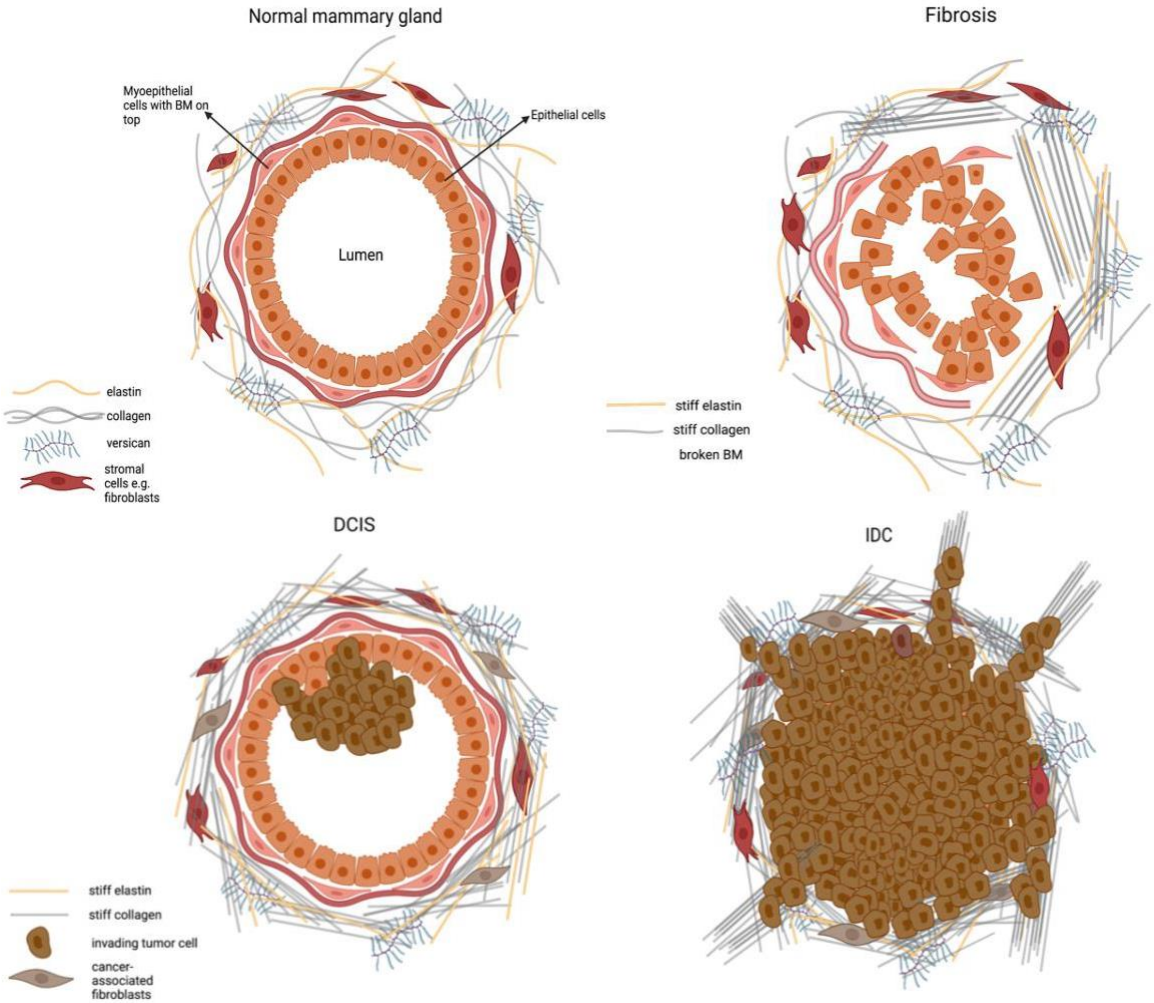


Figure 17: ECM dynamics in fibrosis and breast cancer

Schematic diagrams created with BioRender.com. Adapted from Kaushik et al. [52].

Chapter 8: Conclusion

Co-relation between breast cancer biology and x-ray scatter signals has been successfully outlined in this work. Two main goals were worked upon. First, normal breast tissue and malignant/benign tumors were modeled based on their WAXS signals. The tissue samples were assumed to consist of varying amounts of fat, stroma and cells. μ_s data for breast tissue samples were obtained from literature and that for stroma was an estimate based on histology results by Kidane et *al*. The fractional volumes of each of the three components were calculated using SVD method. The model predictions are promising and give an insight into how cancer interacts with its microenvironment. Second, to have a better understanding of stroma, a stroma model was introduced in which collagen, elastin and versican are considered to be the main components of stroma. Using a ratio of 8:1:1 for amount of collagen, elastin and versican present in stroma, scatter signal for it was produced using the IAM model. The IAM data agrees with the stroma curve (estimated in former studies) at higher end of momentum transfer. The new model is however, not very sensitive to proportion of the three stromal components.

Future work involves collecting experimental WAXS data for cancer cells and breast stroma samples and looking for equivalent biosynthetic materials that can have applications in regenerative medicine and cancer therapy.

Bibliography

[1] "Breast cancer statistics," Canadian Cancer Society, 05 2022. [Online]. Available: https://cancer.ca/en/cancer-information/cancer-types/breast/statistics. [Accessed 05 2022].

[2] C. C. Ontario, "Ontario Breast Screening Program," Cancer Care Ontario, [Online]. Available: https://www.cancercareontario.ca/en/cancer-care-ontario/programs/screening-programs/ontario-breast-obsp. [Accessed May 2022].

[3] R. LeClair, "Modeling Breast Cancer Tissue via Analysis of WAXS signatures," in *World Congress on Medical Physics and Biomedical Engineering*, Toronto, 2015.

[4] G. Kidane, R. D. Speller, G. J. Royle and A. M. Hanby, "X-ray scatter signatures for normal and neoplastic breast tissues," *Physics in Medicine and Biology,* vol. 44, pp. 1791-1802, 1999.

[5] A. Chaparian et al., "The optimization of an energy-dispersive x-ray diffraction system for potential clinical application," *Appl. Radiat. Isot.,* vol. 68, pp. 2237-2245, 2010.

[6] W. Elshemey et al., "The diagnostic capability of x-ray scattering parameters for the characterization of breast cancer," *Med. Phys.,* vol. 37, pp. 4257-4265, 2010.

[7] S. Pani et al., "Characterization of breast tissue using energy-dispersive x-ray diffraction computed tomography," *Appl. Radiat. Isot.,* vol. 68, pp. 1980-1987, 2010.

[8] V. Changizi, M. A. Oghabian, R. Speller, S. Sarkar and A. A. Kheradmand, "Application of Small Angle X-ray Scattering (SAXS) for Differentiation between Normal and Cancerous Breast," *International Journal of Medical Sciences,* vol. 2, no. 3, pp. 118-121, 2005.

[9] O. Oliveira et al., "Identification of neoplasias of breast tissues using a powder diffractometer," *J. Radiat. Res.*, vol. 49, pp. 527-532, 2008.

[10] J. Griffiths et al., "Correlation of energy dispersive diffraction signatures and microCT of small breast tissue samples with pathological analysis," *Phys. Med. Biol.*, vol. 52, pp. 6151-6164, 2007.

[11] E. Ryan and M. Farquharson, "Breast tissue classification using x-ray scattering measurements and multivariate data analysis," *Phys. Med. Biol.*, vol. 52, pp. 6679-6696, 2007.

[12] D. Cunha et al., "X-ray scattering profiles of some normaland malignant human breast tissues," *X-Ray Spectrom*, vol. 35, pp. 370-374, 2006.

[13] R. Leclair, M. Boileau and Y. Wang, "A semianalytic model to extract differential linear scattering coefficients of breast tissue from energy dispersive x-ray diffraction measurements," *Med. Phys.*, vol. 33, pp. 959-967, 2006.

[14] G. Harding, J. Kosanetzky, "Status and outlook of coherent-x-ray scatter imaging,' *J. Opt. Soc.* Am. A 4, 933-944 (1987)

[15] S. Sidhu, K. K. Siu, G. Falzon, S. A. Hart, J. G. Fox and R. A. Lewis, "Mapping structural changes in breast tissue disease using x-ray scattering," *The International Journal of Medical Physics Research and Practice*, vol. 36, no. 7, pp. 3211-3217, 2009.

[16] A. Conceicao, M. Antoniassi and M. Poletti, "Analysis of breast cancer by small angle x-ray scattering (SAXS)," *Analyst*, vol. 134, no. 1077-1082, 2009.

[17] G. Falzon et al., "Wavelet-bsed feature extraction applied to small angle x-ray scattering patterns from breast tissue: a tool for differentiating between tissue types," *Phys. Med. Biol.*, vol. 51, pp. 2465-2477, 2006.

[18] P. Fernandez-Nogueira, M. Mancino, G. Fuster, P. Bragado, M. Prats de Puig, P. Gascon, F. Javier Casado and N. Carbo, "Breast Mammographic Density: Stromal Implications on Breast Cancer Detection and Therapy," *Journal of Clinical Medicine,* vol. 9, no. 3, p. 776, 2020.

[19] A. Round et al., "A preliminary study of breast cancer diagnosis using laboratory based small angle x-ray scattering," *Phys. Med. Biol.,* vol. 50, pp. 4159-4168, 2005.

[20] C. Castro et al., "Coherent scattering characteristics of normal and pathological breast tissues," *Radiat. Phys. Chem.,* vol. 71, pp. 649-651, 2004.

[21] M. Fernandez Martinez, "X-rayscattering and diffraction enhanced imaging studies of in vitro breast tissues," University of Helsinki, Finland, 2006.

[22] A. Aranda-Gutierrez and H. M. Diaz-Perez, "Histology,Mammary Glands," in *StatPearls*, Treasure Island (FL), StatPearls Publishing, 2021.

[23] H. Combrink, "Molecular Screening of the South African Indian Population for BRCA1 and BRCA2 Using High Resolution Melting Analysis," Researchgate, 2016.

[24] Z. Zhu, J. Argen, S. Mannisto, P. Pietinen, M. Eskelinen, K. Syrjanen and M. Uusitupa, "Fatty acid composition of breast adipose tissue in breast cancer patients and in patients with benign breast disease," *Nutrition and Cancer,* vol. 24, no. 2, pp. 151-160, 1995.

[25] S. J. Fletcher, P. A. Sacca, P.-C. Mercedes, A. C. Federico, M. F. Serra, F. E. Santino, C. V. Sasso, C. M. Lopez-Fontana, R. W. Caron, C. J. Calvo and Pistone-, "Human breast adipose

tissue: characterisation of factors thatchange during tumor progression in human breast cancer," *J Exp Clin Cancer Res,* vol. 36, p. 26, 2017.

[26] C. Kothari, C. Diorio and F. Durocher, "The importance of breast adipose tissue in breast cancer," *International Journal of Molecular Sciences,* vol. 21, no. 16, 2020.

[27] S. Gallick, "Histology OLM 4.0," Stephen Gallick, 2013. [Online]. Available: http://stevegallik.org/histologyolm_Ch03_P01.html. [Accessed May 2022].

[28] Z. Werb and P. Lu, "The role of stroma in tumor development," *The Cancer Journal,* vol. 214, no. 4, pp. 250-253, 2015.

[29] "Benign Breast Disease," Cleveland Medical clinic, 2022. [Online]. Available: https://my.clevelandclinic.org/health/articles/6270-benign-breast-disease. [Accessed 2022].

[30] T. A. C. S. m. a. e. c. team, "Fibrocystic changes in the breast," American Cancer Society, 25 January 2022. [Online]. Available: https://www.cancer.org/cancer/breast-cancer/non-cancerous-breast-conditions/fibrosis-and-simple-cysts-in-the-breast.html. [Accessed 7 June 2022].

[31] H. Knipe, "Simple breast cyst," Radiopaedia, 2017. [Online]. Available: https://radiopaedia.org/cases/52358. [Accessed 07 June 2022].

[32] J. M. Park, E. A. J. R. Franken, M. Garg, L. L. Fajardo and L. Niklason, "Breast Tomosynthesis: Present considerations and future applications," *Radiographics,* vol. 27, pp. S231-40, 2007.

[33] M. Ajmal, M. Khan and K. Van Fossen, Breast Fibroadenoma, FL: Statpearls Publishing, 2022.

[34] Z. Li, H. Gao, M. Xu, Y. Wu and D. Gao, "Breast papillary lesions diagnosed and treated using ultrasound-guided vacuum assisted excision," *BMC Surgery*, vol. 20, no. 204, 2020.

[35] A. Ashraf, "Intraductal papilloma of breast," Radiopaedia, 2020. [Online]. Available: https://radiopaedia.org/cases/85487. [Accessed June 2022].

[36] J. Luveta, R. M. Parks, D. M. Heery, K.-L. Cheung and S. J. Johnston, "Invasive Lobular Breast Cancer as a Distinct Disease: Implications for Therapeutic Strategy," *Oncology and Therapy*, vol. 8, no. 1, pp. 1-11, 2019.

[37] M. Colleoni, N. Rotmensz, P. Maisonneuve, M. G. Mastropasqua, A. Luimi, P. Veronesi, M. Intra, E. Montagna, G. Cancello, A. Cardillo, M. Mazza, G. Perri, M. Iorfida, G. Pruneri and G. Viale, "Outcome of special types of luminal breast cancer," *Ann. Oncol.*, vol. 23, no. 6, pp. 1428-36, 2012.

[38] G. Kruger, "Invasive lobular breast carcinoma case study," Radiopaedia, 18 April 2017. [Online]. Available: https://radiopaedia.org/cases/18583. [Accessed 8 June 2022].

[39] H. E. Johns and J. R. Cunningham, "Linear Attenuation Coefficient," in *The Physics of Radiology*, Illinois, Charles C Thomas , 1983, p. 137.

[40] N. McDonald, "MSc Thesis: Fat Subtraction Protocol for Wide-Angle X-ray Scatter Analtysis of Breast Biopsies," Laurentian University, Sudbury, 2016.

[41] F. H. Attix, Introduction to Radiologicl Physics and Radiation Dosimetry, Weinheim: Wiley-Vch Verlag GmbH & Co. KGaA, 2004.

[42] S. Evans, D. Bradley, D. Dance, J. Bateman and C. Jones, "Measurement of small-angle photon scattering for some breast tissues and tissue substitute materials," *Physics in Medicine and Biology*, vol. 36, no. 1, pp. 7-18, 1991.

[43] R. Speller and G. Royle, "Tissue characterisation using low angle x-ray scattering," *Journal of X-ray science and technology*, vol. 3, pp. 77-84, 1992.

[44] R. Lewis et al., "Breast cancer diagnosis using scattered x-rays," *J. Synchrotron Radiat.*, vol. 7, pp. 348-352, 2000.

[45] D. Cunha et al., "X-rayscattering profiles of some normal and malignant human breast tissues," *X-ray Spectrom.*, vol. 35, pp. 370-374, 2006.

[46] M. E. Poletti, O. D. Goncalves and I. Mazzaro, "X ray scattering from human breast tissues and breast equivalent materials," *Physics in Medicine and Biology*, no. 47, pp. 47-63, 2002.

[47] R. J. Leclair and P. C. Johns, "A semianalytic model to investigate the potential applications of x-ray scatter imaging," *Med. Phys.*, vol. 25, no. 6, pp. 1008-1020, 1998.

[48] R. J. LeClair, A. Ferreira, N. McDonald, C. Laamanen and R. Y. Tang, "Model predictions for the wide-angle x-ray scatter signals of healthy and malignant benign duct biopsies," *Journal of Medical Imaging*, vol. 2, no. 4, pp. 043502-1-11, 2015.

[49] J. Kozanetzky et al., "X-ray diffraction measurements of some plastic materials and body tissues," *Med. Phys.*, vol. 14, pp. 526-532, 1987.

[50] J. R. Goldblum, Rosai and Ackerman's Surgical Pathology, Eleventh Edition, 2018.

[51] H. Press, S. Teukolsky, W. Vetterling and B. Flannery, Numerical Recipes in C++, 2nd edition, New York: Cambridge University Press, 2003.

[52] S. Kaushik, M. W. Pickup and V. M. Weaver, "From transformation to metastasis: deconstructing the extracellular matrix in the breast," Cancer and Metastasis, vol. 35, pp. 667, 2016.

[53] Z. I. Khamis, Z. J. Sahab and Q.-X. A. Sang, "Active Roles of Tumor Stroma in Breast Cancer Metastasis," *International Journal of Breast Cancer,* vol. 2012, pp. 574025-10, 2012.

[54] F. Sotoodehnejadnematalahi and B. Burke, "Structure,Functionand Regultion of Versican:The Most Abundant Type of Proteoglycan in the Extracellular Matrix," *Acta Medica Irania,* vol. 51, no. 11, pp. 740-750, 2013.

[55] E. T. Goddard, R. C. Hill, A. Barrett, C. Betts, Q. Guo, O. Maller, V. F. Borges, K. C. Hansen and P. Schedin, "Quantative extracellular matrix proteomics to study mammary and liver tissue microenvironments," *Int J Biochem Cell Biol.,* vol. 81, pp. 223-232, 2016.

[56] C. Frantz, K. M. Stewart and V. M. Weaver, "The extracellular matrix at a glance," *Journal of Cell Science,* vol. 123, pp. 4195-4200, 2010.

[57] P. P. Provenzano, D. R. Inman, K. W. Eliceiri, J. G. Knittel, L. Yan, C. T. Rueden, J. G. White and P. J. Keely, "Collagen density promotes mammary tumor initiation and progression," *BMC Medicine,* vol. 6, no. 11, 2008.

[58] A. Tarakanova, S.-W. Chang and M. J. Buehler, Handbook of Nanomaterial Properties chapter 28, Heidelberg: Springer- Verlag Berlin Heidelberg, 2014.

[59] S. Salesse, L. Odoul, L. Chazee, C. Garbar, L. Duca, L. Martiny, R. Mahmoudi and L. Debelle, "Elastin molecular aging promoted MDA-MB-231 breast cancer cell invasiveness," *FEBS Open Bio,* vol. 8, no. 9, pp. 1395-1404, 2018.

[60] A. D. Theocharis, S. S. Skandalis, T. Neill, H. A. B. Multhaupt, M. Hubo, H. Frey, S. Gopal, Gomes, N. Afratis, H. C. Lim, J. R. Couchman, J. Filmus, R. D. Sanderson, L. Schaefer and I, "Insights into the key roles of proteoglycans in breast cancer biology and translational medicine," Biochim Biophys Acta, vol. 1855, no. 2, pp. 276-300, 2015.

[61]S. Islam and H. Watanabe, "Versican: A Dynamic Regulator of theExtracellular Matrix," Journal of Histochemistry and Cytochemistry, vol. 68, no. 11, pp. 763-775, 2020.

[62] W. W. Du, L. Fang, X. Yang, W. Sheng, B. L. Yang, A. Seth, Y. Zhang, B. B. Yang and A. J. Yee, "The Role of Versican in Modulating Breast Cancer Cell Self-renewal," *Molecular Cancer Reserch AACR*, vol. 11, no. 5, pp. 443-55, 2013.

[63] C. de Lima, S. J. J. de Arimateados, A. Nazario and Y. Michelacci, "Changes in glycosaminoglycans and proteoglycans for normal breast and fibroadenoma during menstrual cycle," Biochem Biophys Acta, vol. 1820, pp. 1009-1019, 2012.

[64] M. Badaoui, C. Mimsy-Julienne, C. Saby, L. V. Gulick, M. Paretti, P. Jeannesson, H. Morjani and H. Ouadid-Ahidouch, "Collagen type 1 promotes survival of human breast cancer cells by overexpressing Kv10.1 potassium and Orai1 calcium channels through DDR1-dependent pathway," *Oncotarget*, vol. 9, no. 37, pp. 24653-24671, 2017.

[65] NCBI, "COL1A1 collagen type I alpha 1 chain [Homo sapiens (human)]," PubMed, [Online]. Available: https://www.ncbi.nlm.nih.gov/gene/1277. [Accessed 1 2021 07].

[66] B. Alberts, A. Johnson, J. Lewis, D. Morgan, M. Raff, K. Roberts and P. Walter, Molecular Bioloy of The Cell, New York: Garland Science, Taylor and Francis group, 2014.

[67] M. Naso, D. Zimmermann and R. Iozzo, "Characterization of the complete genomic structure of human versican gene and functional analysis of its promotes," J. Biol Chem, vol. 269, no. 52, pp. 32999-3008, 1994.

[68] R. Iozzo, M. Naso, L. Cannizzaro, J. Wasmuth and J. McPherson, "Mapping of the versican proteoglycan gene (CSPG2) to the long arm of human chromosome 5 (5q12-5q14)," *Genomics,* vol. 14, no. 4, pp. 845-851, 1992.

[69] M. Dours-Zimmermann and D. Zimmermann, "A novel glycosaminoglycan attachment domain identified in two alternative splice variants of human versican.," *J Biol Chem,* vol. 269, no. 52, pp. 32992-8, 1994.

[70] B. Brooksby, B. W. Pogue, S. Jiang, H. Dehghani, S. Srinivasan, C. Kogel, T. D. Tosteson, J. Weaver, S. P. Poplack and K. D. Paulsen, "Imaging breast adipose and fibroglandular tissue molecular signatures by using hybrid MRI-guided near-infrared spectral tomography," *PNAS,* vol. 103, no. 23, 2005.

[71] J. Hubbell, W. Veigele, E. Briggs, R. Brown, D. Cromer and R. Howerton, "Atomic form factors, incoherent scattering functions, and photon scattering cross sections," J. Phys. Chem. Ref. Data, vol. 4, pp. 471-538, 1977.

[72] J. Herrera, C. A. Henke and P. B. Bitterman, "Extracellular matrix as a driver of progressive fibrosis," *The Journal of Clinical Investigation,* vol. 128, no. 1, pp. 45-53, 2018.

[73] T. N. Wight and S. Potter-Perigo, "The extracellular matrix: an active or passive player in fibrosis?," *American Journal of Physiology,* vol. 301, no. 6, 2011.