Università degli Studi di Genova

Scuola di Scienze Mediche e Farmaceutiche

Corso di Laurea Magistrale in Medicina e Chirurgia

Dipartimento di Medicina Interna

Sessione di Laurea del 27 Giugno 2022

Plasma proteomics differentiates patients with early breast cancer from healthy matched controls: an exploratory analysis of the RENOVATE trial

Relatore:

Prof. Gabriele Zoppoli

Correlatore:

Dott. Francesco Ravera

Candidata:

Alessandra Sara Sanna

Anno Accademico 2021-2022

SUMMARY

1. BACKGROUND AND RATIONALE 3

1.1 Breast Cancer: Epidemiology 3

1.2 Histological Classification 3

1.3 Molecular Classification 4

1.4 Current screening, diagnosis methods and main limitations ... 6

1.5 Liquid Biopsy 11

1.6 Study rationale and scopes 16

2. MATERIALS AND METHODS 18

2.1 Study Design 18

2.2 Patients 18

2.3 Sample collection, processing and storage 19

3. RESULTS 25

3.1 Demographics 25

3.2 The concentration of several biologically meaningful proteins significantly differs between patients affected by early BC and matched healthy controls. 26

4. DISCUSSION 28

5. CONCLUSION 30

6. REFERENCES 31

7. ACKNOWLEDGEMENTS 37

1. BACKGROUND AND RATIONALE

1.1 Breast Cancer: Epidemiology

Breast cancer (BC) is the most frequently diagnosed cancer in women, in particular its incidence is currently 11.7% with an estimated 2.3 million new cases per year, surpassing lung cancer[1]. BC is the fifth leading cause of cancer death worldwide, with a mortality rate of 6.9% [1].

In terms of prevalence in Italy, 800.000 women have been diagnosed with breast cancer, the majority of whom are over the age of 75 [2].

In Italy, the incidence is slightly increasing (+ 0.3% per year) [2] but, despite being the leading cause of death from cancer in women, mortality is decreasing significantly (-0.8% per year) [2] .

This reduction in mortality is attributable to both the improvement of screening programs, which allow for BC early diagnosis and intervention, and the advancements in the therapeutic field.

BC is a complex disease, with several histological and molecular subtypes. The diverse histological and molecular features are associated with peculiar clinical characteristics, with a relevant impact on patients' prognosis, risk of recurrence, and response to treatment.

1.2 Histological Classification

The last World Health Organization's histological classification of breast tumors has been published in 2019.

BC histology is classified as follows:

- Ductal carcinoma in situ;
- Lobular carcinoma in situ;
- Invasive ductal carcinoma (ductal breast cancer);
- Invasive lobular carcinoma;

- Medullary carcinoma;
- Mucinous (colloid) carcinoma;
- Tubular carcinoma;
- Papillary carcinoma;
- Metaplastic breast cancer;
- Phyllodes tumors;
- Mammary Paget disease;
- Inflammatory breast cancer.

Adenocarcinomas represent more than 95% of breast malignancies [3] . In particular, 50% to 80% of newly diagnosed cases are invasive ductal carcinoma; the rest of the cases being classified as invasive lobular carcinoma or special types [4] .

BC histological classification is prognostically relevant[5]. For example, invasive lobular carcinoma is histologically characterised by small non-cohesive cells[6] due to the lack of expression of E-cadherin for alterations in its CDH1 gene[7] . This feature causes the tumor to spread more easily to the gastrointestinal tract, peritoneum and ovary [6]. Therefore, despite often having features associated with good prognosis, such as, low to intermediate grade, low Ki-67 expression, positive estrogen receptor (ER) expression and absence of HER2 protein overexpression[6], several researches suggest that overall long-term outcomes of invasive lobular carcinoma may have inferior long-term outcomes compared to those for stage-matched invasive ductal carcinoma [6].

1.3 Molecular Classification

Clinical parameters as tumor size, lymph node involvement, presence of metastasis and histological classification, however, are not sufficient to characterize BC behavior [8].

The IHC classification is the most useful in clinical practice for assessing tumor aggressiveness, risk of relapse, response to therapy and it also guides treatment decisions.

Biological sub-type also emerged as an important predictor of survival [9].

Currently, the molecular classification of breast cancer includes four groups as specified in Table 1.

Intrinsic Subtypes (GEP)	IHC Classification (St Gallen)
Luminal A	«Luminal A» ER and/or PR positive HER2 negative Ki-67<14%
Luminal B	«Luminal B (HER2 negative)» ER and/or PR positive Any Ki-67 HER2 overexpressed or amplified
HER2 enriched	«HER2 positive (non luminal)» HER2 overexpressed or amplified ER and PR absent
Basal-like	«Triple negative» ER and PR absent HER2 negative

Table 1: BC Molecular Classification. GEP, gene expression profiling; IHC, immuno-histochemical.

Luminal A tumors typically have a better prognosis compared to the other subtypes. Luminal B BC, compared to luminal A BC, has higher proliferation rate and lower expression of progesterone receptors. Also, Luminal A typically presents fewer mutations and chromosomal copy-number alterations[10].

HER2 oncogene, located on chromosome 17, codes for HER2, which is a tyrosine kinase receptor. In HER2-enriched BC either amplifications of HER2 gene or overexpression of HER2 protein can occur[11]. In both cases, HER2 amplifications are associated with an adverse prognosis. Although HER2-enriched tumors are more aggressive than HER2- BC, the development of target treatments based on monoclonal antibodies (e.g. trastuzumab) or tyrosine kinase inhibitors (e.g. lapatinib) has critically improved the prognosis of patients with such alteration[11]. HER-2-positive tumors also benefit from anthracyclines [12] and have shown a significantly better response to taxanes than HER2-negative tumors [13].

Triple-negative breast cancer is a subtype of BC characterised by the lack of expression of estrogen receptor and progesterone receptor and the absence of HER2 amplification. It accounts for about 15-20% of all BC diagnosed each year. Triple-negative breast cancers are more aggressive than the other subtypes and are more frequent in women of black race and under 40 years of age[14].

Actually, there is an 80% overlap between 'triple-negative' and intrinsic 'basal-like' subtype. 'Triple negative' also includes some special histological types such as adenoid cystic carcinoma [15].

Although HER2+/ER and basal-like tumors have a worse prognosis than Luminal types, they have a better response to systemic treatments either in the adjuvant or the neoadjuvant setting with higher rates of pathologic complete response to neoadjuvant therapy [16].

1.4 Current screening, diagnosis methods and main limitations

Early detection of BC is one of the most effective strategies to reduce its overall mortality, according to the paradigm of secondary prevention.

Current screening protocols aimed at the early diagnosis of BC are based on mammography and/or ultrasound. In particular, in Italy, mammography is recommended biannually in asymptomatic women aged between 50 and 69 years old and annually in women aged between 45 and 49 [2].

In addition, the extension of screening campaigns to women aged 70-74, is suggested[2].

The detection of BC in its early stages, besides allowing for a less invasive surgery, has a relevant impact on cancer-related mortality. The reduction in mortality in the age group between 50 and 69 is, indeed, 23% for all women invited for mammography and 40% for women who have joined the screening [2].

Digital mammography has an overall accuracy of ~89% [17,18] in detecting BC, with a sensitivity of ~97%, a specificity of ~65%, a positive predictive value of ~89%, and a negative predictive value of ~91%.

Ultrasound is typically performed in younger women, in case of dense breast as a complement for mammography, or in case of self-detection of breast or axillary nodules. Ultrasound has a sensitivity of ~80% and a specificity of ~88% [19].

For women bearing high risk of BC, the method of choice is contrast-enhanced magnetic resonance imaging (MRI) to be performed once a year.

The class of "high risk" is defined by the following conditions:

- BRCA1 or BRCA2 mutation;
- 20-25% lifetime risk according to common risk prediction criteria;
- Li-Fraumeni, Cowden or Bannayan-Riley-Ruvalcaba's Syndromes;
- Previous thoracic radiotherapy in 10-30 aged women.

Other indications for MRI include:

- Preoperative staging of newly diagnosed BC;
- Evaluation of the response to NACT;
- Differential diagnosis of pericicatricial lesions;
- CUP syndrome;
- Equivocal results at mammography or ultrasonography;
- Clinical or instrumental suspect in women with breast prosthesis.

The detection of asymptomatic breast lesions is extremely frequent in the overall female population. In most cases these incidental findings are fibroadenomas, which are the most common benign breast lesions. Their prevalence is 25% and they usually occur between 15 and 35 years of age. On mammography, fibroadenomas appear as well-circumscribed masses, with or without popcorn-like calcifications, and do not require further examination. Percutaneous biopsy is recommended for histological diagnostic confirmation in the following cases [20]:

- Suspect findings at mammography or ultrasound (BI-RADS 4 or 5);
- Firm mass in a patient with positive family history or BRCA mutation;

- Clinical or sonographic evidence of tendency to grow;
- New palpable mass in post-menopausal women;

BI-RADS Classification

Breast Imaging-Reporting And Data System (BI-RADS) represents the main classification system aimed at the standardization of BC risk assessment for radiologically detected breast lesions. Proposed by the American College of Radiology in 1986, it has been revised in 2003, 2006, and 2013 [21]. According to BI-RADS, each report must include the specification of the clinical setting (i.e. screening or diagnostic, in case of self-detection of a symptomatic lesion), the description of breast density (fatty, scattered, heterogeneously dense or extremely dense) and possible breast lesions in the form of masses, calcifications, asymmetries or architectural distortions.

In case of detection of a breast lesion, its features should be thoroughly described, mentioning shape, margins, and density.

Irregular shape, spiculated margins and high density are typical features of malignant lesions, while regular shape, round margins, and low density are associated with a benign phenotype. Calcifications can have several radiological correlates.

Amorphous, fine pleomorphic, or fine-linear branching calcifications represent suspect findings, while rod-like, popcorn, coarse, vascular, and milk of calcium calcifications are associated with a benign phenotype. The distribution of the calcifications is relevant as well, possibly being diffuse, regional, grouped, linear, or segmental. The clinical manifestations associated with malignant lesions include skin or nipple retraction, skin or trabecular thickening, and axillary adenopathy.

BI-RADS final assessment includes seven categories displayed in Table 2 [21]:

BI-RADS	Description	Managment	Likelihood of Cancer
0	Adequate assestment could not be performed	Needs additional evaluation	N/A
1	Normal findings/Negative	Routine annual screening	Essentially 0%
2	Benign lesion	Routine annual screening	Essentially 0%
3	Probably Benign lesion	Short interval follow-up (6 months)	<2%
4	Suspicious	Perform biopsy	
	4a		2-9%
	4b		10-49%
	4c		50-94%
5	Highly suggestive of malignancy	Perform biopsy	>95%
6	Known biopsy-proven		Proven Malignancy

Table 2: BI-RADS Classification.

Two techniques can be used for such purpose namely Core Needle Biopsy (CNB) and Vacuum-Assisted Breast Biopsy (VABB).

CNB represents the traditional method for breast biopsies and is recommended in case of mass-like lesions evident at the ultrasound, while VABB is a more recent technique recommended for clusters of microcalcifications, not often observable at the ultrasound.

Other indications for VABB include palpable and non-palpable nodular lesions, BI-RADS 3 and 4A. In particular, for lesions <5 mm VABB should be the technique of choice as CNB may give false negatives[22]. VABB seems to have higher accuracy than CNB with a lower rate of histological underestimation [23]. VABB main complications include pain and bleeding, both during and after the procedure [24].

While lesions with BI-RADS 4c or 5 present a high risk of being malignant (50-95% and > 95% respectively), BI-RADS 4a and 4b lesions are mostly benign, with a relative risk of BC of 2-10% and 10-50% respectively[25]. BC is therefore confirmed only in a limited fraction of biopsied lesions - one case out of three in referral centres[26] - the other cases being, in fact, benign lesions (BL). Lesions smaller than 2 cm (cT1) are especially difficult to differentiate, and the appropriateness of risk assessment strongly depends on the radiologist's experience.

Limitations of current diagnostic approaches

Currently, mammography and ultrasound represent the methods of choice for BC diagnosis and screening, despite their suboptimal accuracy, about 85% and 92% respectively [27,28].

In particular, overdiagnosis is one of the main limitations of such approach, with only 30% of women undergoing breast biopsy being actually diagnosed with BC[29].

Breast biopsy bestows a high burden on both patients, in terms of pain and anxiety for the long turnaround time, and the national healthcare system (NHS), in terms of facilities and person-time.

Radiologically suspect breast lesions, classified as BI-RADS 4 or BI-RADS 5, often result in histologically benign lesion, with a relevant discordance rate with the pathologist's assessment [30] . In such cases, breast biopsy is typically repeated and will lead to BC diagnosis in up to 30% of cases[30], with all the side effects and discomfort that each biopsy entails.

BC is confirmed only in 1/3 of biopsied lesions at best, the other being benign lesions. Patients with radiologically suspect benign lesions do not benefit from breast biopsy, bearing lesions with no relevant evolutionary potential to BC.

Moreover, conventional biopsy draws only a small portion of the radiologically detected mass that may not reflect the characteristics of the entire tumor, given the intra-tumor heterogeneity. Another element to consider is inter-tumor heterogeneity since the same cancer can have multiple localizations [31]. This heterogeneity and potential evolution over time may lead to the execution of multiple biopsies, with the aforementioned discomfort and possible side effects.

Given its intrinsic invasiveness, breast biopsy cannot be applied as a standard procedure for cancer molecular follow-up during the administration of systemic therapy in stage IV patients.

1.5 Liquid Biopsy

In oncology, the concept of liquid biopsy refers to the assessment of circulating analytes derived from tumor mass or microenvironment in different biofluids such as blood, urine, saliva, lymph, cerebrospinal fluid, ascites, or seminal fluid [32].

Peripheral blood, in particular, represents the most studied source for biomarker assessment. The advantages of liquid biopsy over conventional tissue biopsy are numerous [32].

First of all, sample collection, is non-invasive or minimally invasive, with lower discomfort and risk of complications compared to traditional biopsies and lower costs for the NHS. A potentially complete tumor profile can be obtained, possibly capturing intra- and inter-tumor heterogeneity with a relatively small amount of blood (typically 6–10 mL of blood) [33].

Moreover, the use of liquid biopsy may allow for a real time assessment of tumor response to therapy, perfectioning current protocols for longitudinal monitoring of stage IV patients [34]. The lack of standardised and validated procedures for the assessment of circulating biomarkers is a considerable limitation for an effective transition of liquid biopsy to clinical practice [35].

Circulating Biomarkers

Circulationg biomarers include circulating tumor cells (CTCs), cell free nucleic acids (DNA, mRNA, micro-RNA, non-coding RNA), proteins, exosomes and tumor "educated platelets" (TEPs) [36] .

cfDNA

Cell-free DNA (cfDNA) is one of the most promising biomarkers in translational oncology. cfDNA is released passively from apoptotic or necrotic cells or actively secreted through vesicles such as exosomes and microvesicles from living cells. In

healthy subjects, cfDNA is mainly derived from apoptosis of blood nucleated cells at low levels and may increase in inflammatory conditions, after physical exercise [37].
Patients with malignant tumors have higher levels of cfDNA than patients with benign lesions, especially in more advanced stages of disease [38].

ctDNA

Circulating tumor DNA (ctDNA) is the fraction of cfDNA derived from apoptotic or necrotic tumor cells, living tumor cells, or CTCs. In the latter two cases an active secretion of ctDNA from extracellular vesicles, such as exosomes and prostasomes, occurs. ctDNA has peculiar molecular features associated with malignancy, such as mutations, copy number alterations, or methylation changes, that can be assessed for clinical purposes.
ctDNA amount depends on several tumor characteristics, such as tumor burden, vascularization, and localization. Among the total cfDNA, ctDNA fraction can range from 0.013% to more than 90% [39].
ctDNA half-life in the bloodstream typically ranges from 16 min to 2.5 h [40].

Cell-free methylated DNA

DNA methylation plays a fundamental role in cellular development and homeostasis with specific cell/tissue methylation profiles. In cancer, a global hypomethylation associated with hypermethylation of CpG-rich regions typically occurs, disrupting cell cycle dynamics and initiating or promoting carcinogenesis. Altered cfDNA methylation patterns can be exploited for clinical purposes, such as early diagnosis of cancer. Tumor-derived cfDNA typically represents <1% of total cfDNA, especially in early stage cancer. Moreover, cfDNA concentration is often low, amounting to ~10 ng/ml, making the identification of cfDNA methylation patterns somehow difficult [41].
Cell-free methylated DNA immunoprecipitation and high-throughput sequencing (cfMeDIP-seq) allows for the assessment of methylation changes with a small quantity of input DNA (1-10 ng), possibly identifying the alterations of cfDNA methylation that occur in early stage cancer, discriminating between different types of tumor [41]. Such technique has already shown outstanding, albeit preliminary, results in the early detection of renal cell carcinoma [42] and intracranial tumors [43].

Proteomics

Proteomics is the study of the entire set of proteins in human cells, tissues, or biofluids[44]. Proteomics includes the assessment of protein concentration, structure, and function.

In oncology, proteomics can play a role for several clinical purposes, such as cancer early diagnosis and follow-up. In proteomics, a qualitative and quantitative profile of numerous cancer-related proteins can be simultaneously assessed. For example, as cancer progresses, changes in the circulating proteome with the alteration of the concentration of several proteins in tissues, blood, or other body fluids often occur. Most of the FDA-approved biomarkers are protein-based biomarkers [32].

Examples include prostate-specific antigen, carbohydrate antigen 125, and carbohydrate antigen 19-9 used in prostate, ovarian, and pancreatic cancer respectively. These markers are mainly used in the assessment of response to therapy and in follow-up to detect relapses but, cannot be used for cancer early diagnosis due to their low specificity and sensitivity [45].

Liquid biopsy potential applications

Liquid biopsy can be used for several clinical purposes, such as early diagnosis, treatment selection and monitoring, and early detection of relapse [46].

Early cancer detection

An accurate assessment of circulating biomarkers released by tumor bulk or microenvironment would allow for the development of methods aimed at the non-invasive diagnosis of cancer. The assessment of tumor-specific mutations in cfDNA, a method considered promising for such purpose, proved not to be accurate enough for an effective transition in clinical practice. Such limitation is due to the limited quantity of ctDNA released from tumor cells, particularly in early stage cancer, and the presence of non-cancer mutations in cfDNA mainly derived from clonal hematopoiesis. A potential solution to this limitation may rely on the combination of different biomarkers, such as proteins, cfDNA fragmentation, mutations, or cfDNA methylation changes.[47,48,49,50].

In a study by Cheng *et al,* cfDNA methylome assessment by cfMeDIP-seq allows for the prediction of BC up to seven years before clinical presentation [51].

In particular, there has been an increasing consensus among recent studies that cfDNA methylation profiles, possibly combined with other markers, such as circulating tumor proteins, and with imaging, may significantly overcome the accuracy of currently approved methods for cancer early diagnosis.
The PCR-based NGS test CancerSEEK, which combined the assessment of ctDNA with several circulating protein biomarkers, is another example of combinatorial approach, that although achieved suboptimal results.[52]

Treatment selection

To date, tumor tissue assay still represents the gold standard for the detection of oncogenic driver variants of genes, such as epidermal growth factor receptor in non-small cell lung cancer and KRAS in colorectal cancer for its higher sensitivity. However, PCR-based cfDNA assays have shown high specificity (96%) [53].
For such purpose, several NGS-based multigene liquid biopsy assays have been approved in clinical practice. Such assays can identify genomic alterations as single-nucleotide variants, insertions, and deletions. Guardant360 assay, for example, is able to detect mutations in ESR1 and PTEN associated with intrinsic resistance to treatment with aromatase inhibitors plus alpelisib [54] and it is useful in treatment selection.

Monitoring treatment efficacy [46].

Currently, CT or other radiological procedures represent the standard methods used to monitor response to treatment. It is possible that, in the next future, liquid biopsy biomarkers such as CTCs or ctDNA will integrate or replace CT controls [55], allowing for an earlier detection of resistance to systemic therapy, for an early switch to more effective chemotherapeutic agents.
In addition, circulating biomarkers from liquid biopsy can provide information about cancer resistance mechanisms, allowing for the selection of the most appropriate therapy, possibly with lower costs compared to traditional biopsies.

Follow-up

Circulating biomarkers such as ctDNA can be potentially used to detect relapse before its radiological evidence. Typical approaches for early detection of cancer relapse,

conventionally include the identification of tumor specific mutations of DNA extracted from the surgical piece with the construction of probes for their identification in cfDNA during follow-up. Despite the potential, no assay is currently approved in clinical practice for such purpose [46].

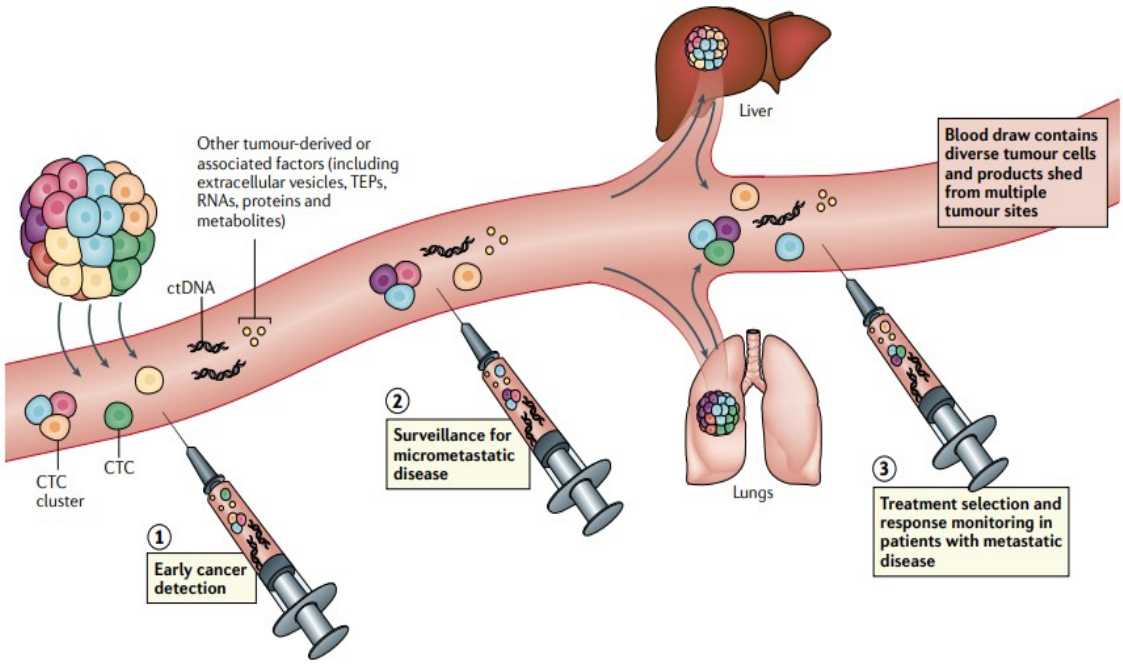

Figure 1: Courtesy of Ignatidis et al. Nature 2021. Clinical applications of liquid biopsy and clinically meaningful biomarkers CTCs, circulating nucleic acids or other tumor-derived materials in the bloodstream[46].

Liquid biopsy actual applications

The potential applications of liquid biopsy are gradually becoming a reality, with an increasing number of tests approved by national regulatory agencies. CellSearch® CTC enumeration platform is the first liquid biopsy test based on the assessment of CTC concentration approved by the Food and Drug Administration (FDA). This test is aimed at the prognostic stratification of patients affected by advanced metastatic breast, colon, or prostate cancer [56]. Few years later, the FDA approved the first ctDNA genetic test to detect epidermal growth factor receptor gene mutations in patients with non–small cell lung cancer, aimed at therapy selection [56].

In 2020, the FDA approved two comprehensive genomic profiling liquid biopsy tests, including Guardant360 CDx, which detects alterations in more than 60 gene, and FoundationOne Liquid CDx, which is able to detect alterations in more than 300 genes

involved in the genesis of solid tumors, besides microsatellite instability and mutations in blood malignancies [56].

Integrated liquid biopsy

Integrated liquid biopsy refers to the combined assessment of several biomarkers or other methods to improve sensitivity and specificity reached with the single methodologies. The integration of multiple techniques within the same sample set has proved to significantly enhance the overall accuracy of the proposed tests, with important implications, concerning the clinical applicability of novel non-invasive methods for clinical purposes. Limitations of each technology may indeed be complemented by strengths of the other, resulting in more accurate classifiers.

There are different types of data integration [57]:

1. Elementary integration, where data of the same type are combined (protein-protein, RNA-RNA, DNA-DNA);
2. Intermediate integration, where two or more different types of biomarkers are integrated (e.g. protein-DNA);
3. Advanced integration, where liquid biopsy is integrated with other kinds of analyses, such as radiology.

1.6 Study rationale and scopes

The combination of circulating biomarkers with radiomics may represent a turning point in the early differential diagnosis of BC, possibly allowing to avoid unnecessary invasive tests, such as breast biopsy, and improving the overall accuracy of current screening protocols.

This work illustrates the preliminary results of the RENOVATE trial, which aims at creating an integrated classifier based on the assessment of several cutting-edge methodologies for the differential diagnosis of suspect breast lesions. Especially, plasma proteomics, cell-free methylated DNA, and radiomics will be assessed on a cohort of patients with suspect breast lesions with indication for breast biopsy.

The final goal of this study is to create a non-invasive classifier for the assessment of suspect breast lesions detected by mammography or ultrasound, that will allow to avoid breast biopsy in patients with benign lesions. The availability of such classifier will

result in a significant reduction of distress for patients bearing radiologically suspect benign breast lesions and lower costs for the National Health System.

Moreover, in case of success, this project will foster the design of prospective trials aimed at testing our classifier as a novel screening tool, possibly enhancing the accuracy of current screening protocols and resulting in a better management of BC patients with significant reduction of their mortality rate.

In this work, we will focus on RENOVATE preliminary proteomics data; we will illustrate the results obtained from plasma proteomics assessed in two exploratory cohorts of patients affected by early BC and matched healthy individuals. Such results will be subsequently integrated with data from cfMeDIP-seq and radiomics performed on the same cohort in the prosecution of the RENOVATE trial.

2. Materials and Methods

2.1 Study Design

Patients evaluated at the Diagnostic Senology Unit with evidence of BIRADS-3/4/5 lesions ≤ 2 cm (radiological T1), were asked to donate 4 peripheral blood tubes for a total of ~35 mL with appropriate preserving solutions and 1 urine sample of ~50 mL. Samples were also collected from a cohort of healthy controls, i.e. healthy women with two consecutive negative mammograms.

San Martino Hospital's Diagnostics Senology Unit is one of the highest-level referral institutions in Italy, serving a population of over 2,000,000 people. Of the 15,000 mammograms performed in a year, 1,500 lead to a radiological suspicion of malignancy, which requires a tissue biopsy for a definite diagnosis.

Clinically-meaningful data were collected as well, including age, body mass index, smoke, consumption of alcohol, menopausal status, age at menarche, and comorbidities. A second blood and urine collection (T1) was performed only in patients diagnosed with BC after breast surgery. Proteomics have been exploratively assessed on samples collected before the diagnostic biopsy from a cohort of patients affected by early BC and from a cohort of matched healthy women. cfMeDIP-seq and radiomics will be assessed on the same cohorts in order to build an exploratory integrated classifier aimed at the differential diagnosis of suspect breast lesions.

2.2 Patients

Patients selected for this study include women who have been found to have suspect breast lesions at screening bilateral mammography screening and scheduled for breast biopsy, and a cohort of healthy controls with two consecutive negative mammography.

Inclusion Criteria	Exclusion Criteria
Written informed consent;	Previous history of cancer;
Women with breast lesions detected by digital bilateral mammography (for study population) or women without detectable breast lesions at the mammography (two consecutive BI-RADS1 mammograms for healthy controls);	Clinical or radiological suspicion of advanced or metastatic cancer at the time of screening;
Age ≥ 18 years and ≤ 75 years;	Known history of active or treated autoimmune or manifest chronic or seasonal and active allergic disorders (with the exception of autoimmune thyroiditis);
Eligible for diagnostic biopsy (tru-cut or VABB) as per normal clinical practice for study population;	History of major trauma or surgery during the 24 weeks before screening;
Ability and willfulness to comply with the protocol requirements.	History of active infectious disease, either chronic or acute but occurring during the 8 weeks before screening;
	History of known acute or chronic cardiac, kidney, or liver disease.

Table 3: Inclusion and Exclusion Criteria

2.3 Sample collection, processing and storage

Two biologists and a research nurse collaborated to collect 2 PAXgeneBlood ccfDNA tube, one BD Vacutainer® K2 EDTA Plus Tube, one Life Technologies Tempus® tube (~35 mL total), and one urine specimen in a sterile container. DICOM files of the diagnostic mammograms have been collected as well (Figure 2).

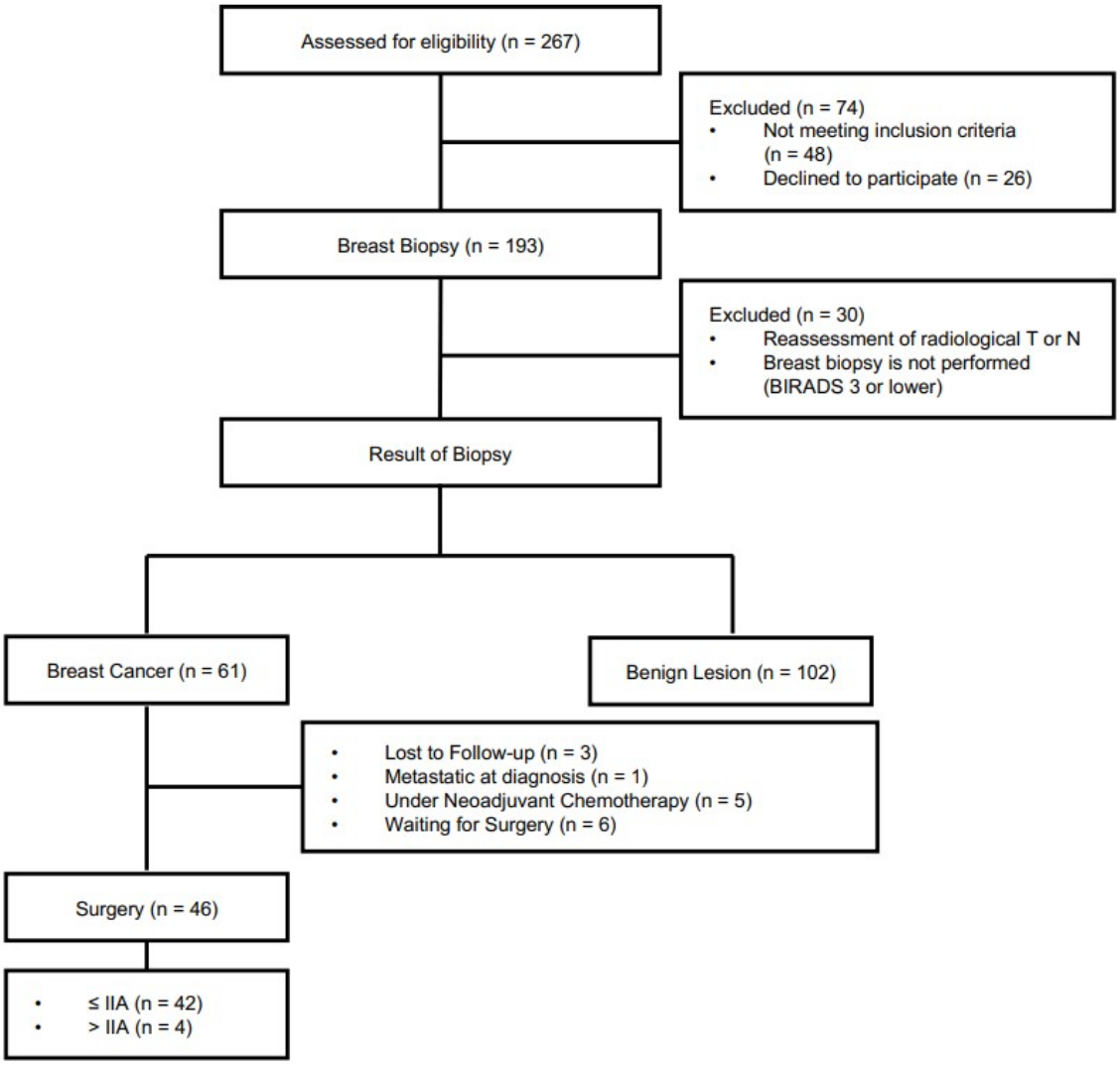

Figure 2: CONSORT-like chart of RENOVATE trial recruitment.

Once the histological diagnosis of BC was confirmed, a second blood and urine sample was collected after surgery. Samples are stored in a dedicated double redundancy freezer (electrical backup and CO2 tank) at -80 °C. Each tube is barcoded to ensure traceability and anonymity, both at the time of collection and during storage.

Blood and urine tubes were processed after less than two hours from collection. Blood samples are processed to extract and store cfDNA from plasma, and to collect proteins, exosomes, peripheral blood mononuclear cells, and total RNA. PAXgene tubes are first centrifuged for 15min at 1900 rcf at room temperature (RT), then the collected plasma is further centrifuged for 10 min at 1900 rcf at RT. EDTA tubes are centrifuged for 15min at 1600 rcf at RT and the collected plasma is further centrifuged at 1900 rcf for 10 min at RT. Plasma is then aliquoted in cryovials. Tempus tubes for total RNA extraction are immediately stored at −80°C. At the time of collection, urine is mixed

with Cell-Free DNA Urine Preserve (Streck) in order to stabilize cfDNA in samples for up to 7 days at a temperature ranging from 6°C to 37°C. Urine is centrifuged at 2680 rcf for 10min, and the supernatant is aliquoted into 15 mL tubes and stored at −80°C until cfDNA extraction. Samples are stored at −80°C in a dedicated, Eppendorf CryoCube F740hi ULT Freezer, three compartments.

2.4 Proteomics analysis

A hundred and fifty µL of plasma were aliquoted from EDTA tubes and sent to SomaLogic® for proteomic analysis. SomaLogic® has developed a new, highly multiplexed proteomic assay (SOMAscan ™) for the relative measurement of 7596 blood proteins based on SOMAmers (Figure 3). SOMAmers are a novel type of aptamer consisting of single-stranded DNA molecules which can bind specific proteins. This assay includes a unique 40-nucleotide sequence tag and a fluorescent label that allows for protein identification and quantification in high-density microarrays.

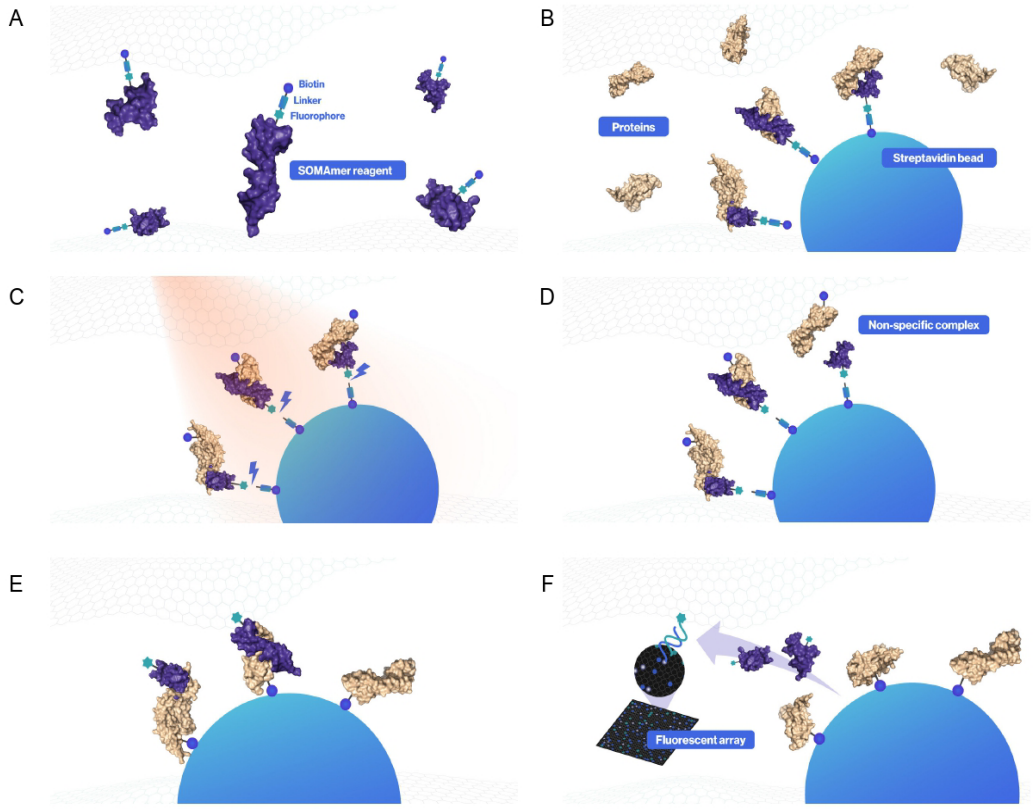

Figure 3: Figure 3 SOMAscan ™ White Paper.

Steps of protein detection using SOMAscan ™. SOMAmer reagents are synthesized with a fluorophore, photocleavable linker, and Biotin (A). SOMAmer reagents bound to streptavidin beads are used to capture proteins from a complex mixture of proteins (B). UV light breaks the photocleavable linker, releasing complexes back into solution (C). Non-specific complexes dissociate while specific complexes remain bound (D). Biotinylated proteins (and bound SOMAmer reagents) are captured on new streptavidin beads (E). SOMAmer reagents are released from the complexes by denaturing the proteins. Fluorophores are measured after hybridization to complementary sequences on a microarray chip. The fluorescence intensity detected on the microarray is related to the amount of available epitope in the original sample (F).

2.5 Statistical analyses of proteomics results

SomaLogic® adopts an internal calibration method, which is proprietary to the Company, by which scalar dilutions of the assessed sample are combined in a known concentration with spike-in controls. Dilution curves are then generated by the Company, and the provided results are a centered and scaled matrix of protein expression values from the desired biofluid. Such matrix, comprising 7,596 protein expression vectors, is parsed into R/bioConductor using the limma package, and analyzed by general linear modeling adjusting for fixed variables and adjustment ones. The resulting log fold changes are ranked, and nominal p-values are corrected by

multiple adjustment (Benjamini-Hochberg, i.e. False Discovery Rate correction). For protein set enrichment analyses and overrepresentation analyses, log fold changes are ranked and scores from known gene signatures are generated using the mSigDB database and the ClusterProfiler package. Heatmaps are generated by hierarchical clustering using the Euclidean distance methods and the Pearson correlation measure. Other statistical tests are used as needed.

2.6 Other proposed analyses

Methylome profiling of cfDNA

cfDNA methylation will be assessed by cfMeDIPseq. cfMeDIP-seq will be performed to identify methylation changes in limited amounts of cfDNA (1-10 ng). We used the QiaCube® technology to extract cfDNA from collected plasma. As with a previous protocol [58], cfMeDIP will be performed following these four steps:

1.cfDNA end repair, A-tailing and adapter ligation;

2.cfMeDIP immunoprecipitation and enrichment using an Ab targeting 5 methylcytosine;

3.Library preparation;

4.High throughput NGS on an Illumina platform for cfMeDNA data.

Radiomics analyses

For the radiomic analysis of the collected mammograms, we will use the preliminary classifier based on digital breast tomosynthesis (DBT) images built in the ASTOUND trial [59]. Radiomics analyses were done on all DBT images within manually selected regions of interest, including all the dense parts of the breast and omitting the fatty portions.

Descriptors for the preliminary classifier were chosen following an initial screening of 104 radiomics characteristics to avoid over-fitting and in accordance with features previously associated with cancer risk in breast parenchymal patterns [59]. An open source software platform for medical image informatics will be used to extract image

features from the same cases and matched controls for whom NGS and proteomics analyses are performed.

HDI classifier

An integrated classifier will be developed by using the ensemble learning approach methodology, integrating data from proteomics, cfMeDIPseq, and radiomics (spazio di troppo).

Ensemble learning combines the predictions of multiple individual classifiers obtained using different techniques, such as random forest, support vector machine, or general linear modelling, to improve generalization power, avoid overfitting, and enhance the strength and reliability of the final result. A weighted-majority voting system implemented in the R environment will be used to integrate the results of proteomics, cfMeDIPseq, and radiomics [60].

3. RESULTS

3.1 Demographics

A total of 193 patients undergoing breast biopsy were recruited. Of these, 61 were diagnosed with BC, 102 with histologically benign lesions, and 30 were excluded. A hundred patients with two consecutive BIRADS-1 mammographies were recruited as well.

A total of 40 patients were selected for the proteomics analyses. Of these, 20 were diagnosed with BC, while 20 were healthy individuals. The median age of BC patients was 56.5 (95% quantile interval: 42.8 - 75.0) while that of healthy controls was 57 (95% quantile interval: 43.325 - 74.050). BMI in the two groups was 22.91 (95% quantile interval: 19.16 - 37.37) and 23.42 (95% quantile interval: 18.25 - 36.17) respectively. In both groups there were 6 pre-menopausal and 14 post-menopausal women. Of the patients diagnosed with BC, 13 were affected by stage IA BC and 5 by stage IIA BC. For two patients pathological staging was not available. Eighteen patients were affected by ductal BC, one by lobular BC, and one by cribriform BC. Among the selected BC patients, 13 were affected by luminal B, six luminal A, one triple negative.

	Patients with BC	Healthy Controls
Age	56.5 (42.8 – 75.0)	57 (43.325 - 74.050)
BMI	22.91 (19.16 – 37.37)	23.42 (18.25 – 36.17)
Pre-menopause	6	6
Post-menopause	14	14

Table 4: Demographic data.

3.2 The concentration of several biologically meaningful proteins significantly differs between patients affected by early BC and matched healthy controls.

Proteomic analysis performed on 20 patients affected by early BC and 20 healthy matched controls showed a significant difference in the concentration of 595 plasma proteins. In particular, the concentration of 399 proteins was significantly higher in patients affected by BC compared to healthy controls. Conversely, the concentration of 196 proteins was significantly higher in healthy individuals compared to BC patients (Figure 4).

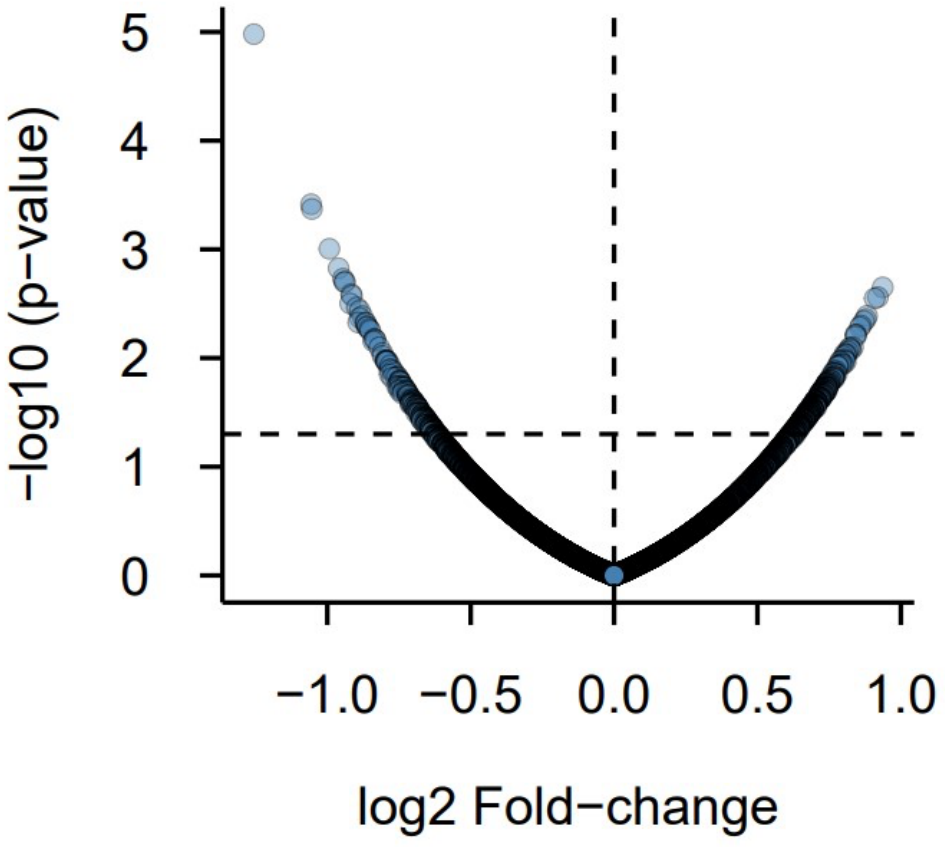

Figure 4: Volcano plot highlighting the difference of protein concentrations assessed with SOMAscan ™ between BC patients and healthy controls.

Protein set enrichment analysis (Figure 5), performed with the aim of evaluating the top proteomic pathways differentially expressed between BC patients and healthy individuals, showed a downregulation of proteomic pathways associated with Epithelial-mesenchymal transition (p = 0.005; FDR= 0.066), IL2 STAT5 signaling (p = 0.02; FDR= 0.10), and KRAS signaling (p = 0.033; FDR = 0.10) in BC patients compared to healthy controls. On the contrary, PI3K-AKT-MTOR signaling pathway was found significantly upregulated in BC patients compared to healthy controls (p = 0.033; FDR=0.10).

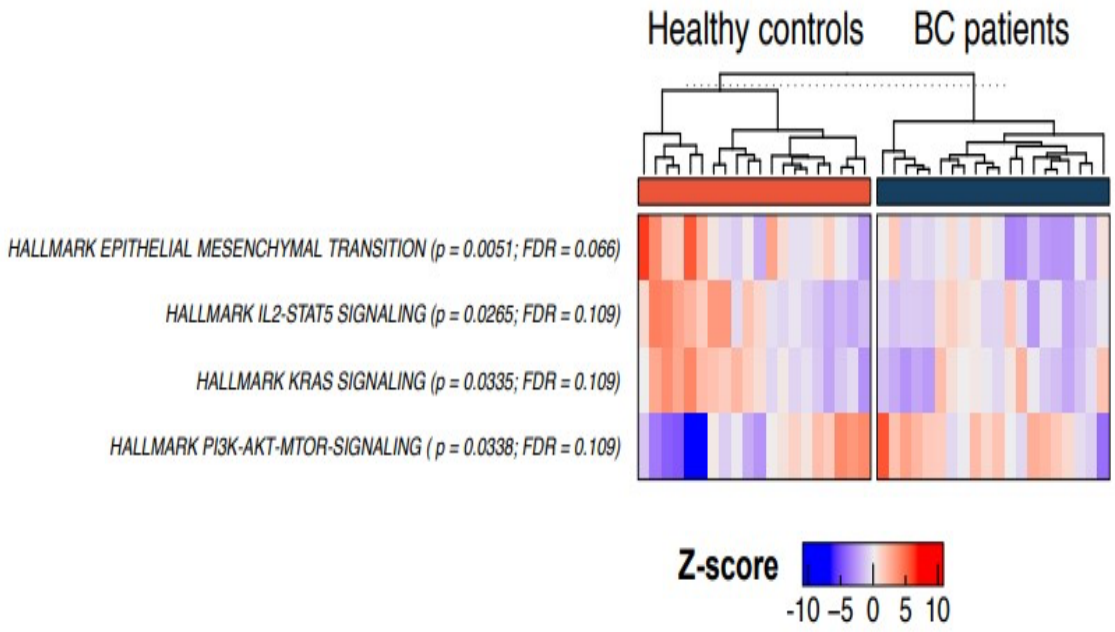

Figure 5: Differential enrichment of the top proteomic pathways differentially expressed between BC patients and healthy controls.

4. DISCUSSION

Plasma proteomics represents one of the several approaches attempted so far for the early detection of cancer. Several studies have been conducted to demonstrate the clinical validity of protein-based biomarkers in clinical practice, thus far with suboptimal results. The assumption is that specific proteins or peptides are secreted or consumed by cancer or cancer microenvironment, resulting in a different concentration of several analytes in the bloodstream of cancer patients compared to healthy individuals [55].

However, the major limitation of proteomics is the lack of sensitivity, specificity, and/or accuracy especially in the detection of early-stage disease. In a study by Kazarian *et al* several markers such as carcinoembryonic antigen, the soluble form of MUC1 protein (carbohydrate antigen 15-3, carbohydrate antigen 27.29), and circulating cytokeratin fragments (TPA, TPS and CYFRA 21-1) have been assessed for such purpose, showing insufficient accuracy for an effective application in clinical practice. Other promising biomarkers, such as carbohydrate antigen 15-3, Carcinoembryonic Antigen, and HER-2 failed to achieve adequate accuracy for the early diagnosis of cancer, but are currently approved in limited settings such as post surgical follow-up [61].

In order to identify early BC, another research group tried to combine the analysis of ten protein biomarkers, namely Carcinoembryonic Antigen, carbohydrate antigen 125, α-fetoprotein, osteopontin, haptoglobin, leptin, prolactin, cancer antigen 19-9 and migration inhibitory factor, achieving a sensitivity of ~ 50% [62].

Analyzing the characteristics of the previous works, it may be observed that some questionable aspects recur. It is worth noting that a small number of proteins is assessed. This obviously reduces the chances of finding combinations of biomarkers that correlate with the presence of malignancy.

It is hardly conceivable, indeed, that the evaluation of a single analyte will achieve sufficient accuracy for diagnostic purposes. The combined assessment of multiple analytes may, however, bring to clinically useful results, possibly allowing for the development of an accurate classifier aimed at the early detection of cancer [63]. Moreover, in many studies, immunoassays are the method of choice, with suboptimal results compared to cutting-edge assays like SOMAscan ™.

In the present work, four protein pathways were found to be significantly different in BC patients compared to healthy controls. Such pathways are all biologically meaningful in cancer initiation or progression.

IL2-STAT5 presents a multifaceted role in BC initiation and progression, with its hypofunction being associated with poorly differentiated carcinomas and its overexpression being associated with well differentiated papillary BC, in a complex interplay with other biologically meaningful pathways such as Wnt, Cav-1 and Notch[64,65].

PI3K-AKT-MTOR pathway is the most aberrant pathway in BC, and is associated with alterations in cell proliferation, angiogenesis, and regulation of apoptosis [66].
EMT-related protein cluster includes proteins involved in cancer progression and dissemination. During EMT cancer cells typically lose cell-cell junctions, apical-basal polarity, epithelial markers, and acquire cell motility, a spindle-cell shape, and mesenchymal markers, with the possibility of migrating to distant site from the primary tumor [67].
While the upregulation of PI3K-AKT-MTOR pathway in BC patients compared to healthy controls is an expectable and consistent finding, the relative downregulation of KRAS signaling pathway and EMT cluster is somewhat controversial. Such observations can be explained by pre- and post- translational modifications of proteins included in such pathways occurring in cancer. Such alterations, which may include changes in mRNA splicing and DNA transcription, phosphorylation, methylation, glycosylation, acylation, oxidation, or ubiquitination may result in a suboptimal biding between altered proteins and SOMAmers, developed on the basis of physiological protein conformation, with an apparent reduction of their concentration in cancer patients [68].

Other hypotheses may include a local and/or distant feedback loop. Proteins pertaining to KRAS pathway and EMT may be sequestered by breast stroma surrounding the tumor, in a local reaction aimed at containing cancer progression, or their secretion in the bloodstream may be inhibited by the presence of downstream peptides in a negative feedback loop.

5. CONCLUSION

Our results, albeit preliminary, point toward the potential role of plasma proteomics in the early differential diagnosis of BC.

The analysis of plasma samples collected from patients with benign lesions will allow us to establish, and possibly validate, the clinical-validity of limited protein sets, leading to the design of prospective trials aimed at the evaluation of their clinical utility, possibly in combination with other biomarkers such as cell-free methylated DNA.

For such purpose, further studies aimed at assessing protein structural modifications are needed to understand their biological relevance and their potential clinical applicability. In the end, it is possible that no method will be accurate enough for clinical transferability alone. It is more likely that the two molecular approaches are investigated together, in order to build an integrated, robust classifier. Thus, limitations of each technology may be complemented by strengths of the other.

6. REFERENCES

1. Sung, H. *et al.* Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA: A Cancer Journal for Clinicians* **71**, 209–249 (2021).
2. *Linee guida NEOPLASIE DELLA MAMMELLA*. (2019).
3. Vinay K, A. K. J. C. N. F. Robbins and Cotran Pathologic Basis of Disease. Eight ed. Elsevier; Lyon, France: . (2010).
4. Henry, N. L. & Cannon-Albright, L. A. Breast cancer histologic subtypes show excess familial clustering. *Cancer* **125**, 3131–3138 (2019).
5. Ellis, 0 *et al. Pathological prognostic factors in breast cancer. 11. Histological type. Relationship with survival in a large study with long-term follow-up*. *Histopathology* vol. 20 (1992).
6. Rakha E & Ellis I. Breast cancer prognostic classification in the molecular era:the role of histological grade. *Breast Cancer Research* (2010).
7. Reed, M. E. M. C., Kutasovic, J. R., Lakhani, S. R. & Simpson, P. T. Invasive lobular carcinoma of the breast: Morphology, biomarkers and 'omics. *Breast Cancer Research* vol. 17 (2015).
8. Masood, S. Breast cancer subtypes: Morphologic and biologic characterization. *Women's Health* **12**, 103–119 (2016).
9. Kunheri, B., Raj, R. v., Vijaykumar, D. K. & Pavithran, K. Impact of St. Gallen surrogate classification for intrinsic breast cancer sub-types on disease features, recurrence, and survival in South Indian patients. *Indian J Cancer* **57**, 49–54 (2020).
10. Prat, A. *et al.* Clinical implications of the intrinsic molecular subtypes of breast cancer. *Breast* **24**, S26–S35 (2015).
11. Krishnamurti, U. & Silverman, J. F. *HER2 in Breast Cancer: A Review and Update*. www.anatomicpathology.com (2014).

12. Muss H & Thor AD. c-erb-2 Expression and responde to adjuvant therapy in women with node-positive early breast cancer. *The New England Journal of Medicine* (1994).

13. Harris, L. N. *et al.* Molecular subtypes of breast cancer in relation to paclitaxel response and outcomes in women with metastatic disease: Results from CALGB 9342. *Breast Cancer Research* **8**, (2006).

14. Li, X. *et al.* Triple-negative breast cancer has worse overall survival and cause-specific survival than non-triple-negative breast cancer. *Breast Cancer Research and Treatment* **161**, 279–287 (2017).

15. Goldhirsch, A. *et al.* Personalizing the treatment of women with early breast cancer: Highlights of the st gallen international expert consensus on the primary therapy of early breast Cancer 2013. *Annals of Oncology* **24**, 2206–2223 (2013).

16. Carey, L. A. *et al.* The triple negative paradox: Primary tumor chemosensitivity of breast cancer subtypes. *Clinical Cancer Research* **13**, 2329–2334 (2007).

17. Lehman, C. D. *et al.* Diagnostic accuracy of digital screening mammography with and without computer-aided detection. *JAMA Internal Medicine* **175**, 1828–1837 (2015).

18. Zeeshan, M., Salam, B., Khalid, Q. S. B., Alam, S. & Sayani, R. Diagnostic Accuracy of Digital Mammography in the Detection of Breast Cancer. *Cureus* (2018) doi:10.7759/cureus.2448.

19. Sood, R. *et al.* Ultrasound for Breast Cancer Detection Globally: A Systematic Review and Meta-Analysis. *J Global Oncol* (2019) doi:10.1200/JGO.19.

20. Stachs, A., Stubert, J., Reimer, T. & Hartmann, S. Benign breast disease in women. *Deutsches Arzteblatt International* **116**, 565–573 (2019).

21. Barazi Hassana & Gunduru Mounika. *Mammography BI RADS Grading*. (2021).

22. Cassano, E. *et al.* Ultrasound-guided vacuum-assisted core breast biopsy: Experience with 406 cases. *Breast Cancer Research and Treatment* **102**, 103–110 (2007).

23. Park, H. L., Chang, S. Y., Huh, J. Y. & Kim, J. Y. Is further diagnostic surgery necessary for the benign papillary lesions that are diagnosed by large volume vacuum assisted breast biopsy? *Journal of Breast Cancer* **13**, 206–211 (2010).

24. Park Hai-Lin & Hong Jisun. Vacuum-assisted breast biopsy for breast cancer. *Gland Surgery* (2014).

25. Magny SJ, S. R. K. al. *Breast Imaging Reporting and Data System.* (2021).

26. Ghosh, K. *et al. Breast Biopsy Utilization A Population-Based Study*.

27. Pereira, R. de O. *et al.* Evaluation of the accuracy of mammography, ultrasound and magnetic resonance imaging in suspect breast lesions. *Clinics* **75**, 1–4 (2020).

28. Sadigh, G., Carlos, R. C., Neal, C. H. & Dwamena, B. A. Accuracy of quantitative ultrasound elastography for differentiation of malignant and benign breast abnormalities: A meta-analysis. *Breast Cancer Research and Treatment* vol. 134 923–931 (2012).

29. Autier, P. & Boniol, M. Mammography screening: A major issue in medicine. *European Journal of Cancer* **90**, 34–62 (2018).

30. Park, V. Y., Kim, E. K., Moon, H. J., Yoon, J. H. & Kim, M. J. Evaluating imaging-pathology concordance and discordance after ultrasound-guided breast biopsy. *Ultrasonography* vol. 37 107–120 (2018).

31. Sprague, B. L. *et al.* National performance benchmarks for modern diagnostic digital mammography: Update from the Breast Cancer Surveillance Consortium. *Radiology* **283**, 59–69 (2017).

32. Ding, Z., Wang, N., Ji, N. & Chen, Z. S. Proteomics technologies for cancer liquid biopsies. *Molecular Cancer* vol. 21 (2022).

33. Oxnard, G. R. *et al.* Noninvasive detection of response and resistance in egfrmutant lung cancer using quantitative next-generation genotyping of cell-free plasma DNA. *Clinical Cancer Research* **20**, 1698–1705 (2014).

34. Bardelli, A. & Pantel, K. Liquid Biopsies, What We Do Not Know (Yet). *Cancer Cell* vol. 31 172–179 (2017).

35. Castro-Giner, F. *et al.* Cancer Diagnosis Using a Liquid Biopsy: Challenges and Expectations. *Diagnostics* **8**, 31 (2018).

36. Poulet, G., Massias, J. & Taly, V. Liquid Biopsy: General Concepts. *Acta Cytologica* vol. 63 449–455 (2019).

37. Maltoni, R. *et al. Cell-free DNA detected by "liquid biopsy" as a potential prognostic biomarker in early breast cancer*. *Oncotarget* vol. 8 www.impactjournals.com/oncotarget/ (2017).

38. González-Masiá, J. A., García-Olmo, D. & García-Olmo, D. C. Circulating nucleic acids in plasma and serum (CNAPS): Applications in oncology. *OncoTargets and Therapy* vol. 6 819–832 (2013).

39. Bettegowda, C. *et al. Detection of Circulating Tumor DNA in Early-and Late-Stage Human Malignancies*. www.ScienceTranslationalMedicine.org.

40. Fernández-Lázaro, D. *et al.* Liquid biopsy as novel tool in precision medicine: Origins, properties, identification and clinical perspective of cancer's biomarkers. *Diagnostics* vol. 10 (2020).

41. Wilson, S. L. *et al.* Sensitive and reproducible cell-free methylome quantification with synthetic spike-in controls. (2021) doi:10.1101/2021.02.12.430289.

42. Nuzzo, P. V. *et al.* Detection of renal cell carcinoma using plasma and urine cell-free DNA methylomes. *Nature Medicine* **26**, 1041–1043 (2020).

43. Nassiri, F. *et al.* Detection and discrimination of intracranial tumors using plasma cell-free DNA methylomes. *Nat Med* **26**, 1044–1047 (2020).

44. Graves, P. R. & Haystead, T. A. J. Molecular Biologist's Guide to Proteomics. *Microbiology and Molecular Biology Reviews* **66**, 39–63 (2002).

45. Landegren, U. & Hammond, M. Cancer diagnostics based on plasma protein biomarkers: hard times but great expectations. *Molecular Oncology* vol. 15 1715–1726 (2021).

46. Ignatiadis, M., Sledge, G. W. & Jeffrey, S. S. Liquid biopsy enters the clinic — implementation issues and future challenges. *Nature Reviews Clinical Oncology* vol. 18 297–312 (2021).

47. Reiter, J. G. *et al.* An analysis of genetic heterogeneity in untreated cancers. *Nature Reviews Cancer* **19**, 639–650 (2019).

48. Bick, A. G. *et al.* Inherited causes of clonal haematopoiesis in 97,691 whole genomes. *Nature* **586**, 763–768 (2020).

49. Shen, S. Y. *et al.* Sensitive tumour detection and classification using plasma cell-free DNA methylomes. *Nature* **563**, 579–583 (2018).

50. Cristiano, S. *et al.* Genome-wide cell-free DNA fragmentation in patients with cancer. *Nature* **570**, 385–389 (2019).

51. Cheng, N. *et al.* Early signatures of breast cancer up to seven years prior to clinical diagnosis in plasma cell-free DNA methylomes. doi:10.21203/rs.3.rs-1203227/v1.

52. Cohen, J. D. *et al.* Detection and localization of surgically resectable cancers with a multi-analyte blood test. *Science (1979)* **359**, 926–930 (2018).

53. Merker, J. D. *et al.* JOURNAL OF CLINICAL ONCOLOGY Circulating Tumor DNA Analysis in Patients With Cancer: American Society of Clinical Oncology and College of American Pathologists Joint Review. *J Clin Oncol* **36**, 1631–1641 (2018).

54. Razavi, P. *et al.* Alterations in PTEN and ESR1 promote clinical resistance to alpelisib plus aromatase inhibitors. *Nature Cancer* **1**, 382–393 (2020).

55. Budd, G. T. *et al.* Circulating tumor cells versus imaging - Predicting overall survival in metastatic breast cancer. *Clinical Cancer Research* **12**, 6403–6409 (2006).

56. Cavallo Jo. The Evolution of Liquid Biopsy in Cancer Care. *The ASCO post* (2021).

57. Qiu, J. *et al.* Refining cancer management using integrated liquid biopsy. *Theranostics* vol. 10 2374–2384 (2020).

58. Shen, S. Y., Burgener, J. M., Bratman, S. v & de Carvalho, D. D. Preparation of cfMeDIP-seq libraries for methylome profiling of plasma cell-free DNA. *Nature Protocols* doi:10.1038/nprot.201.012.

59. Tagliafico, A. S. *et al.* Adjunct screening with tomosynthesis or ultrasound in women with mammography-negative dense breasts: Interim report of a prospective comparative trial. *Journal of Clinical Oncology* **34**, 1882–1888 (2016).

60. Dietterich, T. G. *Machine Learning Research: Four Current Directions*.

61. Kazarian, A. *et al.* Testing breast cancer serum biomarkers for early detection and prognosis in pre-diagnosis samples. *British Journal of Cancer* **116**, 501–508 (2017).

62. Opstal-van Winden, A. W. J. *et al.* A bead-based multiplexed immunoassay to evaluate breast cancer biomarkers for early detection in pre-diagnostic serum. *International Journal of Molecular Sciences* **13**, 13587–13604 (2012).

63. Kim, B. K. *et al.* The multiplex bead array approach to identifying serum biomarkers associated with breast cancer. *Breast Cancer Research* **11**, (2009).

64. Barash, I. Stat5 in breast cancer: Potential oncogenic activity coincides with positive prognosis for the disease. *Carcinogenesis* vol. 33 2320–2325 (2012).

65. Eilon, T. & Barash, I. Different gene-expression profiles for the poorly differentiated carcinoma and the highly differentiated papillary adenocarcinoma in mammary glands support distinct metabolic pathways. *BMC Cancer* **8**, (2008).

66. Xu, F., Na, L., Li, Y. & Chen, L. Roles of the PI3K/AKT/mTOR signalling pathways in neurodegenerative diseases and tumours. *Cell and Bioscience* vol. 10 (2020).

67. Lai, X. *et al.* Epithelial-Mesenchymal Transition and Metabolic Switching in Cancer: Lessons From Somatic Cell Reprogramming. *Frontiers in Cell and Developmental Biology* vol. 8 (2020).

68. Hoffman, M. D., Sniatynski, M. J. & Kast, J. Current approaches for global post-translational modification discovery and mass spectrometric analysis. *Analytica Chimica Acta* vol. 627 50–61 (2008).

7. ACKNOWLEDGEMENTS

Questi sei anni sono stati molto duri per me, e chi mi è stato vicino in questo percorso lo sa bene.

Ci tengo quindi a ringraziare chi mi ha permesso di raggiungere questo obiettivo tanto ambito e che fino a poco tempo fa reputavo irrealizzabile.

Ringrazio il Professor Zoppoli e il Dottor Ravera per avermi accompagnata negli ultimi mesi verso questo traguardo e supportata nella stesura della tesi.

Ringrazio le mie sorelle Alice, Federica, Giulia, Agaliya e Francesca la mia certezza, sempre.

Ringrazio tutta la mia famiglia, i miei amici e compagni di avventura che hanno reso questo percorso un viaggio indimenticabile.

Ringrazio il mio ragazzo, Elia, che mi ha sempre aiutata e sostenuta, ma soprattutto ci ha creduto molto prima che lo facessi io.

Ringrazio più di chiunque altro i miei genitori Sabrina e Gianluca, che sono stati la mia forza in questi sei anni, come in tutta la mia vita.