

MESTRADO EM ONCOLOGIA
ESPECIALIZAÇÃO EM ONCOLOGIA LABORATORIAL

Clinical significance of ARID1A and ANXA1 in HER-2 positive breast cancer

Rita Oliveira

M
2019



Ana Rita Silva Oliveira

Clinical significance of ARID1A and ANXA1 in HER-2 positive breast cancer

Dissertação de Candidatura ao grau de **Mestre em Oncologia** –
Especialização em Oncologia Laboratorial submetida ao Instituto de Ciências
Biomédicas de Abel Salazar da Universidade do Porto

Orientadora: **Professora Doutora Carmen de Lurdes Fonseca Jerónimo**

Professora Associada Convidada com Agregação
Departamento de Patologia e Imunologia Molecular

Instituto de Ciências Biomédicas Abel Salazar - Universidade do Porto

Investigadora Auxiliar e Coordenadora do Grupo de Epigenética e Biologia do
Cancro

Centro de Investigação

Instituto Português de Oncologia do Porto Francisco Gentil, E.P.E

Coorientadora: **Susana Palma de Sousa**

Assistente Graduada

Departamento de Oncologia Médica e Clínica de Patologia Mamária

Instituto Português de Oncologia – Porto

*"Our greatest weakness lies in giving up.
The most certain way to succeed is always to try just one more time."*

Thomas A. Edison



This study was funded by Research Centre of Portuguese Oncology Institute of Porto (CI-IPOP-82-2017 and CI-IPOP-19-2016)

AGRADECIMENTOS

Chegou ao fim uma das etapas mais desafiantes da minha vida, não só a nível profissional como pessoal. A realização deste trabalho não teria sido possível sem a ajuda de algumas pessoas, a quem quero expressar o meu mais profundo agradecimento.

Gostaria de agradecer, em primeiro lugar, à minha orientadora Professora Doutora Carmen Jerónimo por me ter convidado a realizar a minha dissertação no seu grupo e por me ter acompanhado neste percurso. Agradeço também à Dr.^a Susana Palma de Sousa por ter aceitado coorientar a minha dissertação.

Expresso o meu agradecimento ao Professor Doutor Rui Henrique pelas críticas construtivas e partilha de conhecimento nas reuniões de grupo e também ao Professor Doutor Manuel Teixeira, na qualidade de diretor do centro de investigação do IPO do Porto, pela possibilidade de realizar a minha dissertação neste centro de referência.

Ao serviço de anatomia patológica, nomeadamente à Ana Teresa pela realização de todos os cortes usados neste trabalho, à Isa e à Rita Guimarães, pelos cortes dos controlos e, um agradecimento muito especial, à Dr.^a Sara Petronilho que realizou a avaliação de todas as lâminas de imunohistoquímica. Obrigada, Sara, por todas as vezes em que te “chatee” para me seleccionares blocos e delimitares lâminas e, particularmente, pela tua calma e objetividade quando chegava ao pé de ti a “stressar” por algum motivo! Obrigada por tudo.

Ao Serviço de Patologia da Mama, em particular à Dr.^a Filipa pela recolha da informação clínica das doentes incluídas neste estudo.

Ao Serviço de Epidemiologia, nomeadamente ao Engenheiro Luís, pela disponibilidade demonstrada para a realização da análise estatística.

Quero agradecer a todos os membros do Grupo de Epigenética e Biologia do Cancro que foram a minha segunda família neste último ano e meio. Obrigada às meninas “mais velhas” Sara, Sofia, Vânia, Vera e Nair pelo conhecimento transmitido e por me mostrarem o que é trabalhar em investigação. Obrigada ao pessoal da “sala dos mais novinhos” por todos os momentos ali partilhados; desde os “stresses” causados pelo trabalho que corre mal até às pequenas conquistas que juntos celebramos. Todos vocês irão ficar, para sempre, guardados no meu coração.

À Helena, por se ter sempre disponibilizado a ajudar-me, pela leitura da introdução desta dissertação e por ver sempre o lado positivo em tudo! À Dani e à Sandra pela ajuda no desenho de primers de metilação e por nunca deixarem de acreditar que haveríamos de conseguir optimizá-los! Obrigada.

Aos meninos (senhores!) do grupo, Gonças e Zé, obrigada pela vossa boa disposição, partilha de *memes*, comentários sarcásticos e piadas. Sem eles, o nosso dia-a-dia não seria o mesmo!

À “minha” Carina. Obrigada por teres entrado na minha vida. És incrível e um exemplo de força para todos. Desejo-te sempre o melhor que a vida tem para oferecer.

Às minhas colegas de mestrado Cláudia, Mariana, Verita, Teixas (és emprestada!) e Sílvia, obrigada por percorrem este caminho comigo. A vossa presença foi imprescindível para tornar este percurso mais suportável!

À Verita, por toda a ajuda que me deu na fase inicial, por me ter ensinado tantas técnicas indispensáveis para este trabalho e pelo espírito crítico e descontração com que (incrivelmente) lida com todas as situações. Obrigada por seres um exemplo do que é ser uma boa investigadora e, para além disso, uma excelente colega.

À Cláudia, a minha fiel companheira de diversão! Obrigada por todos os momentos que partilhamos. Nunca esquecerei as nossas IHQ e extrações com a tua (não tão) excelente escolha musical, os nossos stresses antes do simpósio, a nossa partilha diária de *memes*, que nos davam incentivo para enfrentarmos mais um dia a rir... basicamente todos os momentos que passamos juntas! Obrigada por estares sempre lá para mim, especialmente nas alturas menos boas.

À Lameirinhas. Por toda a ajuda que me deu ao longo e, particularmente, no final desta etapa, desde a organização de tubos a -80°C, passando pela análise estatística e até à discussão de dúvidas e leitura crítica da tese. Por teres aturado as minhas mensagens incessantes durante a semana e fins de semana e por nunca teres desistido de mim. Por tudo isto te agradeço, mas especialmente pelo teu apoio, que foi muito importante para mim nesta fase final. Palavras não chegam para demonstrar a minha gratidão!

À malta da “*Jantarada!*”, por todos os momentos bem passados fora do lab, que ficarão sempre gravados na minha memória.

Às minhas amigas de sempre, Cat, Elisa e Inês, pela vossa paciência e pelo vosso encorajamento quando mais preciso dele. Obrigada por fazerem parte da minha vida.

À minha família, por todo o apoio que me deram ao longo do meu percurso. Agradeço especialmente aos meus pais por serem os meus pilares, por “estarem lá para mim” quando mais preciso, por me incentivarem a nunca desistir e por tolerarem a minha má disposição, especialmente em alturas críticas. Vocês são um exemplo para mim.

Ao André, por ser o meu porto seguro nos bons e nos maus momentos. Obrigada pelo amor e apoio incondicional que só tu me consegues dar.

RESUMO

Introdução: O Trastuzumab é considerado a terapia *standard* no tratamento das doentes com cancro de mama HER-2 positivo. Independentemente dos benefícios da sua utilização, muitas doentes em estadios precoces apresentam, eventualmente, recorrência da doença, e progressão em menos de um ano. Dado que uma parte significativa das doentes HER-2+ não responde à terapia com Trastuzumab, novos biomarcadores de prognóstico e de resposta são necessários para permitir uma melhor estratificação das doentes. Estudos anteriores reportaram que o gene *ANXA1* contribui para a resistência ao Trastuzumab através da ativação do AKT. A existência de uma associação entre a perda de ARID1A e a sobre expressão de ANXA1 foi também já sugerida. Por estas razões, fomos investigar o valor das proteínas ARID1A e ANXA1 na predição de resposta à terapia com Trastuzumab e no prognóstico de doentes de cancro de mama HER-2+.

Métodos: Neste estudo foram analisadas amostras de 215 doentes com cancro de mama HER-2+. O *status* de metilação do promotor do gene ARID1A foi estudado através de PCR quantitativo específico de metilação. A expressão proteica de ARID1A e de ANXA1 foi avaliada por imunohistoquímica.

Resultados: Na nossa coorte, a hipermetilação do promotor do gene *ARID1A* não parece regular a expressão da proteína. Contrariamente ao esperado, não foi observada uma associação negativa entre a expressão de ARID1A e de ANXA1. Os tumores HER-2+ (não-luminal) exibiam maior expressão de ANXA1 do que os tumores luminal B-*like* (HER-2+). Não foram encontradas diferenças significativas relativamente ao valor preditivo da imunoexpressão destas proteínas na resistência ao Trastuzumab. Todavia, foi encontrada uma associação entre expressão de ANXA1 e menor tempo livre de recorrência, bem como de menor tempo de sobrevivência.

Conclusão: Na globalidade, os nossos resultados sugerem que ANXA1 é um promissor biomarcador de prognóstico em doentes HER-2+. Além disso, a sua capacidade de discriminação entre os subtipos luminal B-*like* (HER-2+) e HER-2+ (não-luminal) poderá ser útil na estratificação das doentes no que diz respeito à estratégia terapêutica.

ABSTRACT

Background: Trastuzumab is considered the standard of care of HER-2+ breast cancer patients. Regardless the benefits of its use, many early stage patients eventually recur, and usually disease progresses within a year. Since a significant part of HER-2+ patients do not respond to Trastuzumab, new prognosis and response predictor biomarkers are warranted to allow a better patient stratification. ANXA1 was previously reported to contribute to Trastuzumab resistance through AKT activation. Furthermore, an association between ARID1A loss and ANXA1 upregulation was also previously suggested. Thereby, we investigated the value of ARID1A and ANXA1 protein levels in Trastuzumab response prediction and clinical outcome in HER-2+ BrC patients.

Methods: In this study, tissue samples from 215 HER-2+ breast cancer patients were examined. Promoter methylation status of ARID1A was evaluated by qMSP and ARID1A and ANXA1 protein expression were assessed by immunohistochemistry.

Results: In our cohort, promoter hypermethylation does not seem to regulate ARID1A's expression. Contrarily to what was expected, no negative association was found between ARID1A and ANXA1 expression. HER-2+ (non-luminal) tumours displayed higher ANXA1 expression than luminal B-like (HER-2+) tumours. Concerning Trastuzumab resistance, ARID1A and ANXA1 proteins did not demonstrate predictive value as biomarkers. Importantly, an association was depicted between ANXA1 expression and early recurrence and shorter survival from the disease.

Conclusion: Overall, our results suggest that ANXA1 may be a useful prognostic marker in HER-2+ patients. Additionally, its ability to discriminate between HER-2+ (non-luminal) and luminal B-like (HER-2+) patients might assist in patient's stratification regarding treatment strategy.

TABLE OF CONTENTS

INTRODUCTION	1
BREAST CANCER	3
Epidemiology	3
Risk Factors	4
Screening and Diagnosis	5
Histological Subtypes	6
Staging	6
Prognosis Factors and Predictive Biomarkers	7
Molecular Subtypes	8
Treatment	9
TRASTUZUMAB	10
EPIGENETICS	12
DNA Methylation	13
Chromatin remodellers	14
AIMS	17
MATERIAL AND METHODS	21
PATIENTS AND SAMPLES COLLECTION	23
DNA EXTRACTION	23
SODIUM BISULPHITE MODIFICATION	23
QUANTITATIVE METHYLATION-SPECIFIC PCR	24
IMMUNOHISTOCHEMISTRY	25
IMMUNOHISTOCHEMISTRY SCORING	25
STATISTICAL ANALYSIS	26
RESULTS	27
CLINICAL AND PATHOLOGICAL DATA	29
ARID1A PROMOTER METHYLATION STATUS	30
RELATIONSHIP BETWEEN ARID1A AND ANXA1	30
ARID1A AND ANXA1 EXPRESSION BY MOLECULAR SUBTYPE	30
RESISTANCE TO TRASTUZUMAB	31
HIGH ARID1A AND ANXA1 EXPRESSION IS ASSOCIATED WITH EARLY RECURRENCE AND SHORTER SURVIVAL	32
DISCUSSION	35
CONCLUSIONS AND FUTURE PERSPECTIVES	41
REFERENCES	45
APPENDIX	55

FIGURE INDEX

Figure 1. Estimated age-standardized breast cancer incidence in women worldwide in 2018. Adapted from ¹	3
Figure 2. Estimated number of new cases of breast cancer in Europe (A) and Portugal (B) in women, in 2018. Adapted from ¹	4
Figure 3. Mechanisms of action of Trastuzumab. The binding of Trastuzumab to the HER-2 receptor leads to recruitment of immune cells (namely natural killers [NK]) to the tumour site, triggering ADCC. It also blocks hetero/homodimerization of the receptor, preventing activation of downstream signalling pathways, consequently suppressing cancer cells growth, proliferation and survival.	11
Figure 4. Composition of SWI/SNF complex. These complexes comprise conserved core subunits (blue) and variant subunits (yellow) (A). Mechanisms of remodelling. These mechanisms include binding of the complex, disruption of histone-DNA contact creating a DNA loop (not shown) that propagates around the nucleosome, generating sites more accessible to DNA binding factors (sliding). It may also happen, although the mechanism is not yet well understood, that adjacent nucleosome ejection may occur subsequently to nucleosome repositioning (B) . Adapted from ⁸¹	15
Figure 5. Relationship between ARID1A and ANXA1. Proportion of tumours with ARID1A high or low intensity and ANXA1 negativity or positivity (A) . Association between ARID1A and ANXA1 [Chi-square: $p=0.183$] (B) . Illustrative images of the different protein intensity scores (C)	30
Figure 6. Distribution of ARID1A and ANXA1 immunoexpression by molecular subtype. Percentage of cases with low and high ARID1A intensity staining score (1+ & 2+ vs 3+) [Chi-square $p=0.749$] (A) . Percentage of cases with or without ANXA1 expression [Chi-square $p<0.001$] (B)	31
Figure 7. ARID1A and ANXA1 as predictors of Trastuzumab resistance. Trastuzumab-resistance free survival curves (Kaplan-Meier with log rank test) of ARID1A (A) and ANXA1 (B) immunoexpression, considering patients that recurred during or within 6 months after Trastuzumab cessation, as resistant.	32
Figure 8. ARID1A and ANXA1 prognostic value. Disease-specific and disease-free survival curves (Kaplan-Meier with log rank test) of ANXA1 (A and C, respectively) and ARID1A (B and D, respectively) immunoexpression.....	33

TABLE INDEX

Table 1. Characterization of Breast Cancer molecular subtypes according to European Society for Medical Oncology (ESMO). Adapted from ¹⁹	8
Table 2. Primers used in qMSP	25
Table 3. Primary antibodies used in IHC and conditions.	25
Table 4. Clinicopathological features of luminal-B like (HER-2+) and HER-2+ (non-luminal) BrC patients enrolled in the study.	29
Table 5. Cox regression models assessing the potential of clinical variables and ANXA1 immunoexpression in the prediction of disease-specific survival.	34
Table 6. Cox regression models assessing the potential of clinical variables and ANXA1 immunoexpression in the prediction of disease-free survival.....	34

LIST OF ABBREVIATIONS

5hmC- 5hmC – 5-hydroxymethyl-cytosine
5mC – 5-methyl-cytosine
ACTB- Actin beta
AJCC- American Joint Committee on Cancer
AKT- Protein kinase B
BMI- Body mass index
BrC- Breast cancer
BRCA1/2- DNA repair associated
CGI's- CpG islands
ChT- Chemotherapy
CpG- Cytosine-phosphate-guanine
DCIS- Ductal carcinoma *in situ*
DFS- Disease free survival
DNA- Deoxyribonucleic acid
DNMT- DNA methyltransferase
DSS- Disease specific survival
EGFR- Epidermal growth factor receptor
EMT- Epithelial- mesenchymal transition
ER- Oestrogen receptor
ERBB2- Erb-B2 receptor tyrosine kinase 2
ESMO- European Society for Medical Oncology
ET- Endocrine therapy
FFPE- Formalin fixed paraffin embedded
FISH- Fluorescence *in situ* hybridization
H&E- Haematoxylin and eosin
HER-2- Human epidermal growth factor receptor 2
IDC- Invasive ductal carcinoma
IGFIR- Insulin-like growth factor 1 receptor
IHC- Immunohistochemistry
ILC- Invasive lobular carcinoma
ISH- *in situ* hibridization
LCIS- Lobular carcinoma *in situ*
MAPK- Mitogen activated protein kinase
MBC- Metastatic breast cancer
MBD- Methyl-binding protein domain

MRI- magnetic resonance imaging
mTOR – Mammalian target of rapamycin
NK- Natural killer
Pi3K- Phosphatidylinositol-3-kinase
PR- Progesterone therapy
PTEN- Phosphatase and tensin homolog
qMSP – Quantitative methylation-specific PCR
RRPA- Reverse phase protein array
RT- Radiotherapy
SSC- Special subtype carcinoma
SWI/SNF- SWItch/Sucrose Non-Fermentable
TCGA- The cancer genome atlas
TET- Ten-eleven translocation
TGFβ- Transforming growth factor beta
TNBC- Triple negative breast cancer
TP53- Tumour protein p53
TRFS- Trastuzumab resistance-free survival
TSG- Tumour suppressor gene

INTRODUCTION

BREAST CANCER

Epidemiology

Breast cancer (BrC) is the most frequent cancer among women, having accounted for approximately 2 million new cases in 2018. It was also the major cause of cancer related death, being responsible for 15% of all deaths, contributing to high morbidity and mortality rates worldwide and constituting a major public health problem¹. Nevertheless, BrC mortality has been declining due to earlier detection and improved treatments².

However, this malignancy's incidence shows a great disparity between high and low income areas, being more frequent in developed regions such as Australia, Western and Northern Europe and Northern America (Figure 1)³.

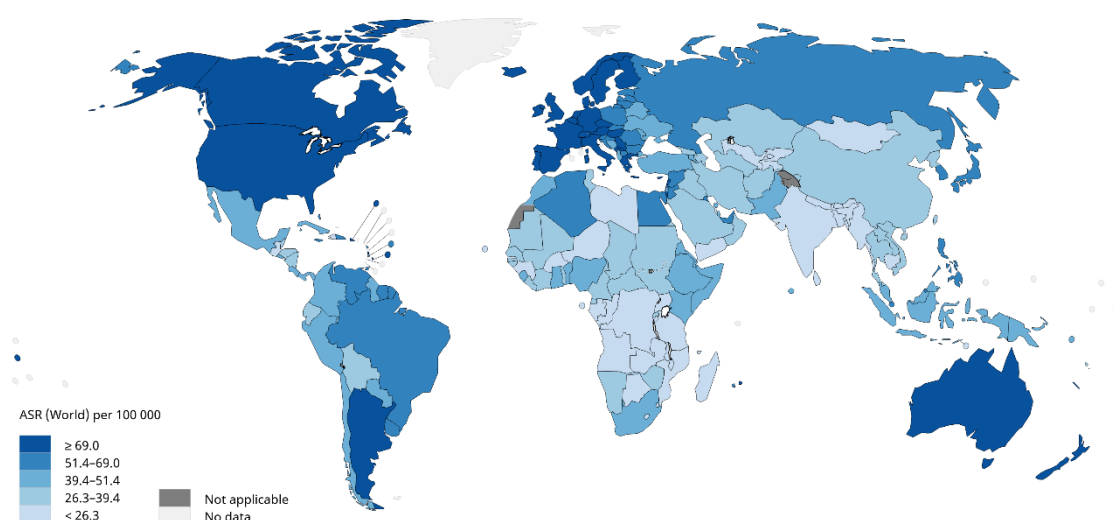


Figure 1. Estimated age-standardized breast cancer incidence in women worldwide in 2018. Adapted from ¹.

The wide geographical variations in BrC incidence are a reflection of differences in risk factors patterns, as well as in availability and access to early detection techniques³. Several factors contribute to these high incidence rates, namely the high prevalence of known risk factors and the implementation of a population based screening^{1,3}.

In Europe, in 2018, BrC was the most common cancer in females accounting for an estimated 26% of all cancers (Figure 2). In Portugal, in the same year, it was also the foremost cancer in women, accounting for an estimated 6974 new cases.

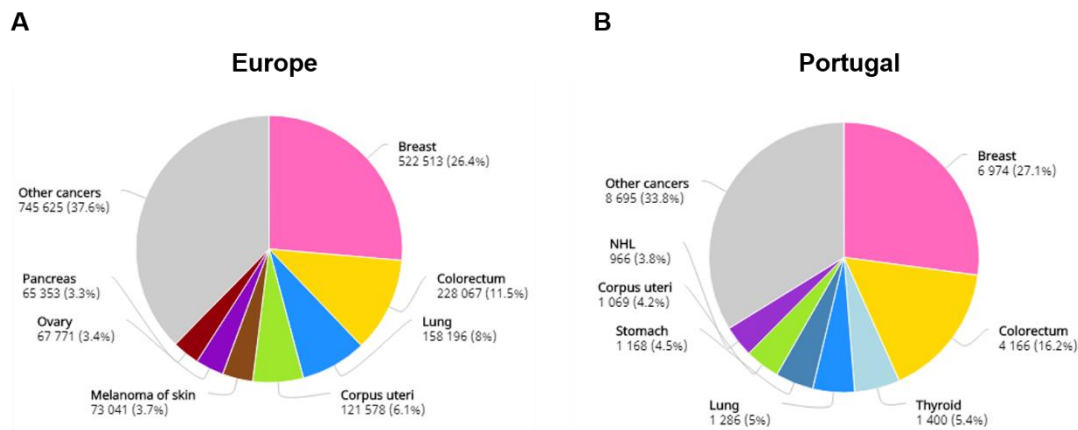


Figure 2. Estimated number of new cases of breast cancer in Europe **(A)** and Portugal **(B)** in women, in 2018. Adapted from¹.

Despite the improvement in early detection, which greatly helped in the prevention of disease progression, BrC still represents a global health care burden, particularly in low and middle-income countries, where the women have limited access to quality treatment^{3,4}. It is imperative to propel awareness into screening and also to improve access to new and better treatments in order to diminish the socioeconomic impact of this disease⁵.

Risk Factors

Several risk factors affect a women's probability of developing BrC, namely age, familial and reproductive history, environmental factors and lifestyle choices (alcohol, diet, obesity and physical activity), among others^{6,7}. Some risk factors are linked to a hereditary component, but most of them are related to the environmental factors to which women are exposed⁸.

Undeniably, age is a very strong factor in BrC incidence, being responsible for a continuously increasing risk beginning from the early reproductive years until the late seventies^{6,9}.

Family history accounts for approximately 10% of BrC incidence in Western countries, and the risk further increases if the affected family member was diagnosed at a young age^{6,10}. Germline mutations in DNA repair associated 1 (*BRCA1*) and 2 (*BRCA2*) genes increase a women's risk in about 40-80%, making these mutations the strongest known predictors¹¹. Mutations in tumour protein p53 (*TP53*) and phosphatase and tensin homolog (*PTEN*) are also involved in familial BrC, however to a lesser extent than the mutations previously mentioned. History of benign breast disease, such as atypical epithelial hyperplasia, is associated with a 4 to 5 times higher probability of developing BrC^{12,13}.

Hormonal factors play an important role in breast carcinogenesis¹⁴. Women with early onset menarche, late age at menopause, late age at first delivery and nulliparous are

at higher risk of developing BrC¹⁴. These risk factors might be explained by their association with oestrogen levels exposition. A higher number of ovulatory cycles and, hence, higher lifetime exposure to this hormone seems to result in a greater risk^{7,10,15}. In contrast, early age at first birth, as well as additional pregnancies and breastfeeding are associated with a decreased risk^{10,16}. The use of oral contraceptives and hormone replacement therapy, being exogenous sources of oestrogen, are also associated with an increased risk¹⁰.

Obesity emerged as an important risk factor due to the “western lifestyle”, namely high intake of saturated fats and lack of physical exercise. However, the relationship between obesity and BrC risk seems to be modulated by the menopausal status. In premenopausal women, a high body mass index (BMI) seems to act as a protective factor whereas in postmenopausal women the opposite is observed¹⁰.

Nevertheless, literature suggests that some lifestyle choices, such as a balanced and healthy diet, engagement in physical activity and low alcohol intake, might reduce a women’s risk of developing BrC^{17, 18}.

Screening and Diagnosis

When detected at early stages, BrC represents a highly curable disease, being one of the reasons why a population-based screening is established nowadays in some countries. In Portugal, as well as in many other European countries, mammography screening is recommended every two years in women aged 50-69 since, at this age group, screening was proved to be most effective in terms of mortality reduction. Additionally, in women with familial breast cancer, annual screening with magnetic resonance imaging (MRI) is also recommended concomitantly or alternating with mammography, starting 10 years younger than the youngest case in the family¹⁹.

Indeed, the population-based screening brought many benefits into the clinical practice, such as allowing tumour early detection, with a concomitant decrease in the comorbidities associated with treatments, and a decrease in mortality rates. Nevertheless, overdiagnosis still represents an important aspect, resulting in the unnecessary treatment of approximately 11% of the women invited to the screening²⁰. Additionally, as radiologic images might be of difficult interpretation, some women are recalled to perform the exam, which causes anxiety and leads to an overall false feeling of security².

BrC diagnosis comprises a clinical examination, with bimanual palpation of the breasts, in combination with imaging, locoregional lymph nodes assessment and confirmation by a pathologist. Biopsy is the gold standard procedure for pathological diagnosis of breast lesions. Fine needle biopsy as well as needle core biopsy can be performed in palpable lesions, however, in non-palpable lesions an ultrasound-guided

biopsy must be performed. If neoadjuvant therapy is required, a core needle biopsy is mandatory to ensure a diagnosis of invasive cancer and to assess immunostaining markers.

Histological Subtypes

BrC is an extremely heterogeneous disease, both morphologically and clinically, exhibiting distinct biological features as well as different clinical outcomes. Currently, the World Health Organization (WHO) recognizes more than twenty distinct entities. The majority of BrC are adenocarcinomas, which means that they arise from mammary parenchymal epithelium and can be divided into two main subcategories: *in situ* and invasive carcinomas²¹.

In situ carcinomas are neoplastic lesions limited to the ducts or lobules and might be further subdivided into ductal carcinoma *in situ* (DCIS) and lobular carcinoma *in situ* (LCIS)²². These are a potentially malignant proliferation of epithelial cells that have not invaded the stroma and, for that reason, do not represent an invasive lesion. Nevertheless, they are precursor lesions to the development of invasive carcinoma, greatly increasing the risk²².

Nevertheless, invasive carcinomas account for most of the cases and comprise a heterogeneous group classified according to cytoarchitectural characteristics, cell type, type of secretion and immunohistochemical profile. This classification allows to distinguish between specific histological types (special subtype carcinoma-SSC) and invasive carcinoma of no special type (NST), the latter, also regarded as invasive ductal carcinoma (IDC), which accounts for approximately 70-80% of all invasive carcinomas^{19,22,23}.

Invasive lobular carcinoma (ILC) is the second most incident subtype, being responsible for approximately 12-15% of invasive cases²⁴. Additionally, tumours that exhibit a combined morphology of NST and SSC, are classified as mixed type tumours²¹.

Staging

The stage of each BrC patient is determined in order to understand the extent of spread of the disease, which helps defining prognosis and the best treatment to be assigned. Currently, the eighth edition of the American Joint Committee on Cancer/ Union for International Cancer Control (AJCC-UICC) TNM staging system is the most accepted and widely used system in BrC staging^{19,25,26}. This system takes in consideration tumour size and extension (T), the spread of cancer to nearby lymph nodes (N) and to other parts of the body (M) (APPENDIX I).

Based on these 3 parameters (T-N-M), patients are staged in one of five categories (0, I, II, III and IV), (APPENDIX II) allowing a better stratification, outcome prediction and an overview of disease burden¹⁹.

Prognosis Factors and Predictive Biomarkers

Despite the usefulness provided by the anatomic TNM staging, patients with the same stage experienced, contrary to the expected, different outcomes which highlighted the need of improving prognosis assessment. Thus, nowadays, some biologic factors are incorporated in the evaluation to accomplish this task²⁵.

Tumour differentiation was the first factor to be included and it reflects how close the tumour morphology is to the morphology of the tissue of origin. The evaluation is accomplished by the incorporation of Nottingham Combined Histologic Grade, which considers three parameters: the grade of architectural differentiation, nuclear pleomorphism and mitotic index.

The oestrogen receptor (ER) is a nuclear transcription factor that stimulates normal breast cells growth and differentiation, upon oestrogen binding. Approximately 75% of BrC overexpress ER, and are characterized by being well differentiated, exhibiting less aggressiveness and having a better outcome, after surgery, than ER negative tumours²³. Age also seems to play a role in ER positivity since older patients display higher rates of tumours positive for this receptor²⁷. ER is a predictor of response to endocrine therapy (ET), including tamoxifen and aromatase inhibitors, although a small fraction of ER negative tumours also benefit from it^{27,28,29}. The progesterone receptor (PR) is also a nuclear transcription factor that contributes to cell proliferation and its expression is regulated by ER expression²⁸. Accordingly, approximately 60% of ER positive tumours are also PR positive, displaying the greatest benefit from ET²⁸.

Erb-B2 receptor tyrosine kinase 2 (ERBB2) [or human epidermal growth factor receptor 2 (HER-2)] is a transmembrane receptor that belongs to the epidermal growth factor receptor (EGFR) family and is involved in the regulation of cell growth, survival and differentiation³⁰. Approximately 20% of BrC patients tumours show gene amplification and overexpress the HER-2 protein, which confers these tumours a worse biologic behaviour as well as clinical aggressiveness^{30,31}. Nevertheless, HER-2 positivity is predictive of favourable response to anti-HER-2 agents, such as Trastuzumab, and to an anthracycline-based chemotherapy (ChT)^{31,32}.

Assessment of the proliferation marker Ki67 may provide additional prognostic and predictive information in clinical practice, especially in ER+/HER-2- BrC. Nevertheless, there are still inconsistencies regarding methodological issues, namely in defining the cut-off value to be used and in results interpretation, which severely limits its utility^{33,34}.

Gene expression profiling techniques have also emerged in clinical practice to aid in systemic treatment decision, when challenging¹⁹. These analysis are able to create multivariate prediction models based on the quantification of ER-related and proliferation genes assessment³⁵. Oncotype DX® (21-gene panel), MammaPrint® (70-gene panel) and

PAM50 (Prosigna) (50-gene panel) are some of such gene panels, however the high cost limit their use in clinical routine³⁶.

Molecular Subtypes

BrC is a very heterogeneous disease, and patients with the same histological type or stage may present different outcomes and respond differently to the same treatments, as previously referred³⁷. Genomic analysis allowed the recognition of 4 BrC groups (luminal A, luminal B, HER-2 enriched and basal-like), also referred to as intrinsic subtypes (Table 1).

Table 1. Characterization of Breast Cancer molecular subtypes according to European Society for Medical Oncology (ESMO). Adapted from ¹⁹.

Intrinsic subtype	Clinicopathological surrogate markers
Luminal A	“Luminal A-like” ER positive HER-2 negative Ki-67 low ¹ PR high
Luminal B	“Luminal B-like (HER-2 negative)” ER positive HER-2 negative Either Ki-67 high or PR low “Luminal B-like (HER-2 positive)” ER positive HER-2 positive Any Ki-67 and any PR
HER-2 enriched	“HER-2 positive (non-luminal)” ER and PR absent HER-2 positive
“Basal-like”	“Triple-negative” ER and PR absent HER-2 negative

¹ Ki-67 scores should be interpreted considering local laboratory values.

Abbreviations: ER, oestrogen receptor; HER2, human epidermal growth factor receptor 2; PgR, progesterone receptor.

These entities display significant differences in terms of incidence, prognosis and treatment^{38,39}.

Luminal BrC, mainly distinguished by the expression of higher levels of ER and luminal epithelial cytokeratins, accounts for approximately 75% of cases^{40,39}. Luminal subtype further subdivides into luminal A and luminal B subtypes. Luminal A: characterised by the expression of higher levels of ER and lower levels of genes related to proliferation comparing with luminal B tumours, being associated with a rather favourable prognosis^{41,42}.

Luminal B gene expression is similar to that of luminal A; however, this subtype displays a decreased ER-related genes' expression and a variable expression of HER-2. Luminal B tumours are associated with worse prognosis than their luminal A counterparts⁴⁰. Nevertheless, both subtypes show good responses to ET, but poor to ChT⁴².

About 20-30% of BrC tumours are HER-2 enriched⁴³, being characterised by a low expression of the luminal clusters and high expression of HER-proliferation genes⁴¹, namely *HER-2*. These tumours' cells display HER-2 protein overexpression which contributes to cellular proliferation, growth and survival, as previously stated^{44,45}. This tyrosine kinase is expressed in normal human breast cells and is necessary for tissue development, but its overexpression leads to an excessive cell proliferation, survival and invasion⁴⁶. The activation of this protein triggers downstream signalling pathways that include the mitogen-activated protein kinase (MAPK) and the phosphatidylinositol-3-kinase/protein kinase B (PI3K/AKT)⁴⁵. Although HER-2 positive tumours have the second poorest prognosis among BrC subtypes, there are available treatments that specifically target HER-2, having a good impact in patients' clinical outcome^{44,46}. Furthermore, these tumours show a greater benefit from chemotherapy treatments with anthracyclines and paclitaxel^{45,47}.

Basal-like tumours are characterized by absence of ER, PR and HER-2 (hence the surrogate definition TNBC) and by the expression of basal epithelial genes. Despite having worse prognosis than the other intrinsic subtypes, they depict a good response to chemotherapy. Furthermore, these tumours also display a high rate of *TP53* mutations, which contributes to their characteristic aggressiveness and poor prognosis^{42,48}.

Currently, ER and PR *status* assessment by immunohistochemistry (IHC) and HER-2 by IHC or *in situ* hybridization (ISH) are part of the clinical routine practice and allow the recognition of 5 surrogate intrinsic subtypes (Table 1). The acknowledgment of these entities aids in prognostication and treatment decision¹⁹.

Treatment

BrC treatment strategy should be determined by a breast specialized multidisciplinary team comprising medical oncologists, surgeons, radiation oncologists, breast radiologists, pathologists and nurses¹⁹. Treatment may involve various modalities, including surgery, ChT, ET, targeted therapies and radiotherapy (RT)^{19,49}.

Surgery is the first option in operable tumours, which accounts for approximately 60-80% of the newly diagnosed cases. Breast-conserving surgery is the preferred approach, being associated with better cosmetic results. Mastectomy is considered in larger tumours, multicentric disease or when required by the patient¹⁹.

Adjuvant RT is highly recommended, reducing the 10-year recurrence risk by 15%. Other adjuvant modalities, such as ChT, ET and targeted therapies may also be

implemented, based on the surrogate intrinsic phenotype, as determined by ER, PR and HER-2 assessment⁴⁹.

Neoadjuvant treatment, including the modalities used in adjuvant therapy (ET, ChT and targeted therapies), may be considered in locally advanced bulky tumours when downstaging is required before surgery. This treatment should begin 2-4 weeks after diagnosis and staging¹⁹.

(Neo)adjuvant systemic modalities should be based on the individuals' risk of relapse, predictive sensitivity to certain types of treatment and the cost-benefit of their use. ChT is more beneficial in ER-negative tumours, being recommended for TNBC, HER-2+ BrC and luminal tumours with high recurrence risk⁴⁹. Gene expression assays, such as Oncotype DX, MammaPrint, Prosigna, may be used to determine the risk of relapse and the benefit of ChT, when in doubt^{19,38}.

Hormonal receptors (HR) positive BrC benefit from ET, and the prescription of which agent to be used should be made based on patients' menopausal status. In premenopausal women, tamoxifen is recommended, whilst in postmenopausal, aromatase inhibitors as letrozole are used.

Anti-HER-2 therapy should be assigned to patients with overexpression of the receptor, meaning all HER-2 enriched and luminal B (HER-2+) BrC. Trastuzumab, a monoclonal antibody, was the first discovered of such therapies, becoming a standard of care in the HER-2+ disease¹⁹. Although a study demonstrated a similar benefit with 9 weeks of administration, the current recommended treatment duration is 1 year^{19,50}. High risk patients (positive lymph nodes or ER-) may also receive the dual blockade with Trastuzumab and Pertuzumab, according to ESMO guidelines¹⁹.

TRASTUZUMAB

As it was formerly stated, HER-2+ BrC cells display HER-2 protein overexpression at the membrane. This protein may form homo or heterodimers with other members of the EGFR family, becoming activated upon phosphorylation of its intracellular domain. This, in turn, leads to activation of multiple downstream signalling pathways, including Pi3K and MAPK, promoting cell growth and proliferation⁴⁶.

Trastuzumab is a monoclonal antibody that binds to HER-2 extracellular domain (Figure 3)⁵¹. Even though it was initially thought to prevent HER-2 dimerization, it is currently known that its effectiveness is also attributed to other mechanisms. Such mechanisms include antibody-dependent cellular cytotoxicity (ADCC), by attracting immune cells to the their binding site and promotion of HER-2 internalization and degradation^{46,52}. Furthermore, Trastuzumab inhibits Pi3K and MAPK signalling pathways activation, promoting cyclin-dependent kinase inhibitor 1B (CDKN1B [or p27^{Kip1}]) induction and concomitant cyclin-

dependent kinase 2 (CDK2) activity reduction, which leads to cell cycle arrest and apoptosis^{46,53}.

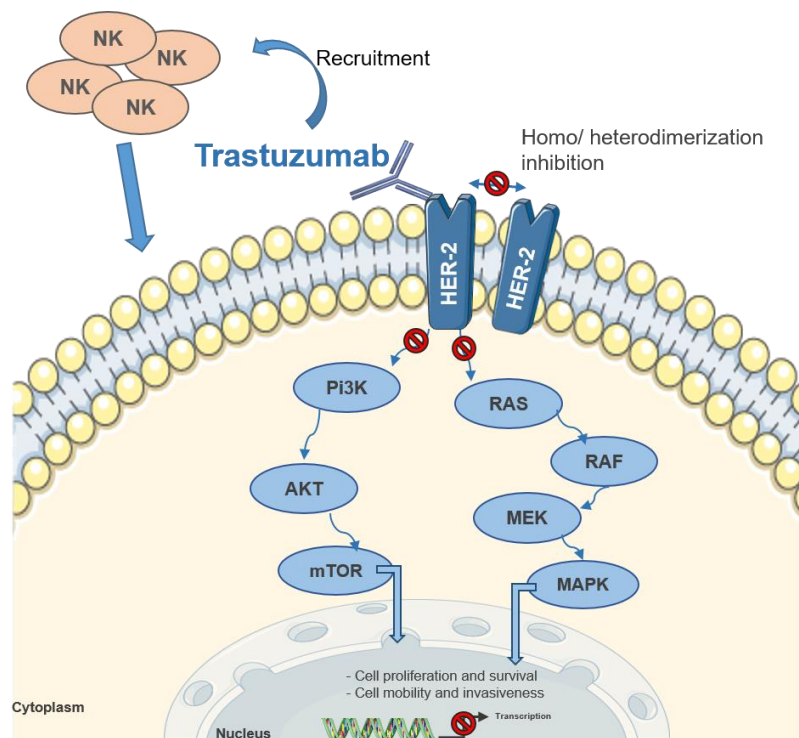


Figure 3. Mechanisms of action of Trastuzumab. The binding of Trastuzumab to the HER-2 receptor leads to recruitment of immune cells (namely natural killers [NK]) to the tumour site, triggering ADCC. It also blocks hetero/homodimerization of the receptor, preventing activation of downstream signalling pathways, consequently suppressing cancer cells growth, proliferation and survival.

It was firstly approved in 1998 for HER-2+ metastatic BrC (MBC) showing an improvement in survival rates, which encouraged researchers to also assess its efficacy and safety in early BrC^{54,55,56}. However, despite the great success brought by Trastuzumab to the management of HER-2+ BrC, approximately 15% of patients with early stage disease develops metastatic disease and 70% of patients with MBC experience progression within a year^{53,57}.

Despite many more studies are warranted in this subject, some have already proposed mechanisms through which cells may be unresponsive to Trastuzumab. One of such mechanisms is altered receptor-antibody interaction; Trastuzumab cannot bind to HER-2 extracellular domain since other molecule is already bound to the protein. One author suggested that the membrane-associated glycoprotein mucin-4 (MUC4) inhibits immune recognition of cancer cells, suppresses apoptosis and activates HER-2, being a contributor to tumour progression and metastasis^{57,58}. Mutations in the *HER-2* gene may also inhibit recognition of the receptor by Trastuzumab, affecting its binding and, therefore, its efficacy⁵⁹. Trastuzumab is able to reduce the activation of HER-2 and respective

downstream pathways. However, even when the antibody is properly bound to the receptor, BrC cells may activate other receptors resulting in the triggering of Pi3K and MAPK signalling pathways or other mitogenic pathways⁵⁹. Furthermore, receptor tyrosine kinases that do not belong to this family, such as the insulin-like growth factor 1 receptor (IGF-IR) may also trigger Pi3K and MAPK pathways, inhibiting Trastuzumab action^{57,60}.

PTEN is also known to downregulate Pi3K and AKT signalling. In fact, loss of its function and concomitant activation of Pi3K, was found in BrC cell lines and further validated in BrC patients' samples⁶¹. Thus, Pi3K upregulation, via mechanisms other than HER-2 activation, is thought to promote AKT activation, which also renders Trastuzumab effectiveness^{62,63}.

In 2015, annexin A1 (*ANXA1*) was reported to associate with Trastuzumab resistance⁶⁴. Many cellular processes have been attributed to this protein, namely calcium signalling, anti-inflammatory effects, receptor mobilization, cell proliferation, and tumour progression. Indeed, *ANXA1* was associated with metastasis promotion in basal-like BrC through regulation of transforming growth factor beta (*TGFβ*) signalling, facilitating an epithelial-mesenchymal transition (EMT)-like switch⁶⁵.

An inverse association between *ANXA1* and adenine thymine-rich interactive domain 1A (*ARID1A*) was proposed by Berns and colleagues⁶⁶ and implicated in Trastuzumab resistance. Their results suggested that loss of *ARID1A* expression is responsible for *ANXA1* upregulation which, in turn, activates Pi3K signalling pathway and, concomitantly, AKT, leading to Trastuzumab resistance⁶⁶. Nevertheless, the authors have only performed *in vitro* assays and validated their findings in datasets from two clinical trials. Additionally, they mainly focused on gene expression evaluation instead of protein, which is more relevant in the understanding of a genes' functional role.

Despite the efforts made to understand the mechanisms through which cancer cells escape the therapeutic action of Trastuzumab, more research on the topic is required since we still lack validated markers of resistance to the antibody. Moreover, in addition to allowing a better patient stratification, the identification of these mechanisms and biomolecules may also be useful for the development of new drugs that may overcome Trastuzumab resistance.

EPIGENETICS

In 1942, Conrad Waddington defined epigenetics as *"the branch of biology which studies the causal interactions between genes and their products which bring the phenotype into being"*^{67,68}. Throughout the years, however, the term evolved and refined, being now accepted as the heritable changes in gene expression without alteration in DNA sequence. Epigenetics are biologically important, taking part in normal cellular processes such as the

embryonic development, X-chromosome inactivation and genetic imprinting, also allowing for the proper regulation of gene expression^{69,70}. Although being mainly considered a genetic disease, epigenetic aberrations play a major role in cancer initiation and development⁷¹.

Epigenetic refers to reversible changes that can lead to altered gene function and neoplastic transformation⁷². The major mechanisms involved in these changes are: DNA methylation, chromatin remodelling, post translational histone modification and non-coding RNAs⁷³.

DNA Methylation

DNA methylation is the most widely studied epigenetic mechanism and the first to be associated with cancer, playing an important role in maintaining genome stability. It consists in the covalent addition of a methyl group to the 5-position carbon of a cytosine present in a cytosine-phosphate-guanine (CpG) dinucleotide, resulting in a 5-methyl-cytosine (5mC)^{73,74}. It is a crucial mechanism that regulates many cellular processes, including transcription, chromatin structure, X chromosome inactivation, chromosomal instability, among others⁷⁰.

Methylation occurs mainly in CpGs within CpG islands (CGIs), which are large clusters of approximately 200-500 base pairs with more than 50% of CpG in content, that are mostly present in genes' promoter regions⁷⁰. This phenomenon can also occur at CpG islands' flanking regions, called CpG shores. These regions are also rich in CpG's, but to a lesser extent than CpG islands, and their methylation is also associated with repression⁷². In normal cells, in transcriptionally active genes, CGIs are unmethylated. On the contrary, genes that do not need to be expressed (in that tissue or particular moment) are characterized by a methylated promoter region⁷⁵.

DNA methylation can induce transcriptional repression through inhibition of binding of transcriptional factors to transcription sites or through the recruitment of transcriptional repressors. Indeed, methylcytosine binding domain (MBD) family members may be activated and, through chromatin remodelling, lead to repression^{69,75}. Nevertheless, mechanisms of transcriptional repression are still a subject of intense investigation.

In cancer cells, both hypo and hypermethylation can occur. A global decrease in methylated CpG content may contribute to activation of silenced genes and genomic instability with proto-oncogenes activation^{69,75}. Conversely, DNA hypermethylation in promoter CpGs, that would otherwise be unmethylated, are associated with transcription repression of tumour suppressor genes (TSGs), which leads to impairment in DNA repair, cell signalling and cell cycle regulation^{72,75}.

DNA methylation is catalysed by DNA methyltransferases (DNMTs), namely DNMT1, DNMT3a and DNMT3b. The latter two enzymes are most commonly associated with *de novo* methylation, being highly expressed during embryonic development and contributing to the establishment of proper patterns of DNA methylation, whereas DNMT1, is responsible for the maintenance of the established DNA methylation patterns^{72,75}. Conversely, ten-eleven-translocation (TET) proteins have the ability to revert the methylation by converting methylated cytosines into hydroxymethylated cytosines (5hmC), which can, then, be excised⁷⁶.

As previously stated, DNA methylation is important in cancer development, as it leads to alteration of chromatin conformation, regulating both the silencing of TSG as well as the activation of oncogenes⁷². Indeed, this epigenetic process is emerging in the field of biomarkers, since many genes' promoters have already been identified as methylated and, therefore, inactivated⁷⁷. Methylation patterns can be created with information regarding these genes' alterations, with the purpose of being used as diagnostic, prognostic and predictive biomarkers⁷⁸. Additionally, since these modifications can be reversed, therapeutic approaches are an appealing emerging field of research⁷⁹.

Chromatin remodellers

The DNA strand, if stretched out, would measure several meters in length. So, the strand must be in a highly condensed conformation in order to fit the nucleus of a cell. This condensed state is achieved through the formation of a complex named nucleosome, in which DNA is tightly packed around proteins called histones⁸⁰. *ARID1A* encodes a subunit of the SWItch/Sucrose Non-Fermentable (SWI/SNF), a nucleosome remodelling complex that uses the energy of ATP to regulate gene expression by altering chromatin conformation (Figure 4)⁸¹.

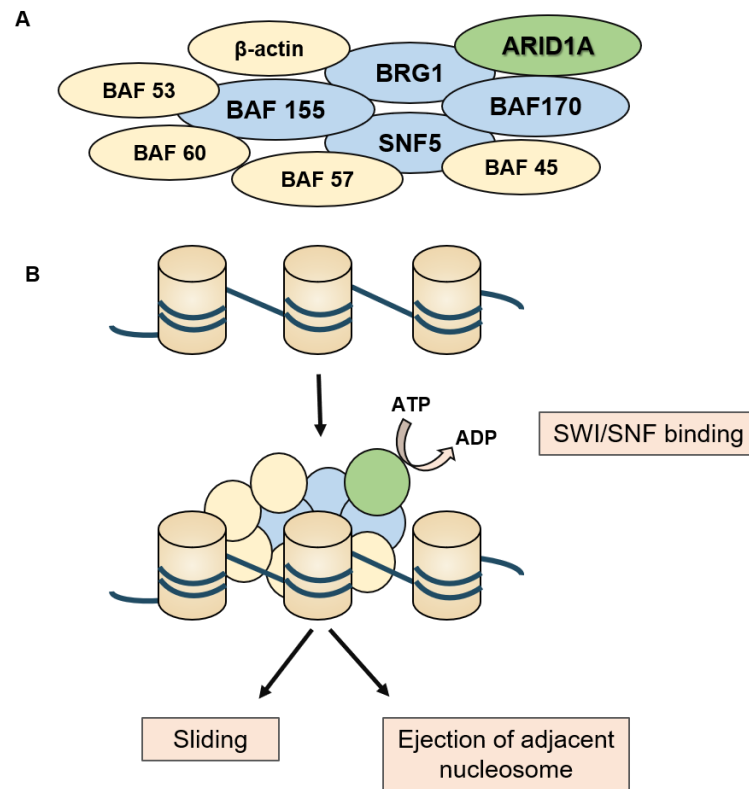


Figure 4. Composition of SWI/SNF complex. These complexes comprise conserved core subunits (blue) and variant subunits (yellow) **(A). Mechanisms of remodelling.** These mechanisms include binding of the complex, disruption of histone-DNA contact creating a DNA loop (not shown) that propagates around the nucleosome, generating sites more accessible to DNA binding factors (sliding). It may also happen, although the mechanism is not yet well understood, that adjacent nucleosome ejection may occur subsequently to nucleosome repositioning **(B)**. Adapted from⁸¹.

ARID1A downregulation might affect the complex function, altering the expression of genes that are under its control by disrupting nucleosome sliding activity and recruitment of coactivators/corepressors^{81,82}. Mutations in *ARID1A*, with concomitant downregulation, have been reported in several cancers such as ovarian clear cell carcinoma⁸³, uterine⁸⁴, endometrial⁸⁵ and gastric cancers⁸⁶, among others. Nevertheless, according to data from the cancer genome atlas (TCGA) and the catalogue of somatic mutations in cancer (COSMIC), *ARID1A* mutations occur with a frequency of 4% in BrC. Additionally, copy number loss was also implicated in *ARID1A* downregulation in BrC, however, to a small extent (13%)⁸⁷. Hence, other mechanisms may account for *ARID1A* downregulation. Indeed, recently, Zhang *et al.*⁸⁸ reported that promoter hypermethylation was responsible for *ARID1A* downregulation in BrC. Despite also having evaluated the contribution of copy number loss and mutations, their results showed that promoter hypermethylation was the main responsible for *ARID1A* reduced mRNA expression levels. Furthermore, H3K27Me3, a transcription repressive histone marker, was significantly overexpressed in samples with low *ARID1A* levels. Indeed, gene promoter hypermethylation is often accompanied by this repressive marker⁶⁹, thus suggesting a role of epigenetic mechanisms in *ARID1A*

downregulation. Since Zhang *et al.*⁸⁸ only evaluated ARID1A mRNA levels, further studies elucidating the role of these mechanisms in protein expression are warranted.

AIMS

BrC is a heterogeneous disease, comprising tumours with different molecular features that are considered for patients' clinical management. Nevertheless, BrC is still the main cause of cancer-related death among women and, so, new biomarkers, able to improve prognosis and disease characterization, are required.

Trastuzumab revolutionized the treatment of HER-2+ BrC, however, many patients do not seem to benefit from it, which dampens its effectiveness. Hence, it is important to find new biomarkers of prognosis and treatment prediction in order to understand which patients are most likely to benefit from Trastuzumab.

ANXA1 regulation by ARID1A had previously been addressed, but not specifically in a cohort of HER-2+ BrC patients. Additionally, the process through which ARID1A exerts such modulation is still unknown. Also, to our knowledge, there are no studies that evaluate ARID1A and ANXA1 immunoexpression with prognosis and, specifically, Trastuzumab resistance in a cohort of HER-2+ BrC treated with a Trastuzumab-based therapy.

Thus, the main goal of this work is to investigate if ARID1A might be regulated by methylation and the value of ANXA1's and ARID1A's protein expression in prognosis and prediction of Trastuzumab resistance in the HER-2+ disease.

Thus, the specific tasks of this work were:

- Assessment of the methylation status of ARID1A promoter in a cohort of HER-2+ BrC FFPE specimens
- Evaluation of ARID1A and ANXA1 protein expression in the same cohort
- Comparison of ARID1A and ANXA1 immunoexpression with clinicopathological parameters, and determination of their prognostic and predictive value

MATERIAL AND METHODS

PATIENTS AND SAMPLES COLLECTION

The study cohort comprises 215 consecutive HER-2+ BrC patients, diagnosed from 2008 to 2013 and treated with a Trastuzumab-based therapy at the Portuguese Oncology Institute of Porto, Portugal (IPO-PORTO). Formalin-fixed paraffin embedded (FFPE) tissue samples analysed were collected before patients' treatment.

Hematoxylin and eosin (H&E) slides were reviewed by an experienced pathologist according to the World Health Organization (WHO) classification⁸⁹. Relevant clinical data was collected from clinical records and displayed in an anonymized data base for analysis purposes.

This study was approved by the institutional review board (Comissão de Ética para a Saúde- CES 125/019) of IPO-PORTO.

DNA EXTRACTION

Nucleic acids were extracted from histological sections (8 µm of thickness) of FFPE blocks using a commercial extraction kit (FFPE RNA/DNA Purification Plus Kit, Norgen Biotek, Thorold, Canada) following manufacturer's instructions.

FFPE were deparaffinized and digested with proteinase K [20mg/mL (NZYTECH, Portugal)] for 15 minutes at 55°C and centrifuged. Digestion buffer and proteinase K were added to the pellets followed by an overnight incubation at 55°C with agitation. In the following day, temperature would be increased to 90°C for 2h. Afterwards, buffer RL and ethanol were added to the tubes and the DNA purification column was mounted. The DNA-containing solution was then added to the column and centrifuged enabling the binding of DNA to the membrane. The membrane-bound DNA was then washed with the provided wash solution and eluted.

Following, DNA purity ratio and concentration were determined using a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA) and then stored at or -20°C.

SODIUM BISULPHITE MODIFICATION

This technique consists on the treatment of denatured DNA with sodium bisulphite, which deaminates cytosine residues converting them into uracil residues. This protocol allows the distinction between methylated and unmethylated cytosine upon quantitative PCR amplification since, following bisulphite treatment, uracil is amplified as thymine whereas methylated cytosines remain cytosines.

One µg of genomic DNA and CpGenome™ Universal Methylated DNA (Millipore, USA) were modified using EZ DNA Methylation-Gold™ Kit (Zymo Research, USA) following manufacturers' instructions.

CT conversion reagent was added to the tubes containing the samples and incubated in the Applied Biosystems Veriti 96-Well Thermal Cycler (Thermo Fisher Scientific, USA) at 98°C for 10 minutes and 64°C for 3 hours. Next, the samples were transferred to the Zymo-Spin™ IC column with 600µL of M-binding buffer. Samples were centrifuged and washed with M-Wash buffer. M-Desulphonation buffer was added to the column, followed by a 20-minute incubation at room temperature. Another centrifugation was performed, and the columns were washed two more times.

Genomic DNA samples and CpGenome™ Universal Methylated DNA were eluted in 60µL or 30µL of sterile distilled water, respectively. The modified DNA was stored at -80°C until further use.

QUANTITATIVE METHYLATION-SPECIFIC PCR

Quantitative real time methylation specific PCR (qMSP) was performed in order to assess the methylation levels in all samples. QMSP is a sensitive and specific method that allows the assessment of a gene's promoter methylation status. Methyl Primer Express Software v1.0 (Applied Biosystems, Foster City, CA, USA) was used to predict regions enriched in CpG dinucleotides (CpG islands) in the *ARID1A*'s region previously described by Zhang *et al.*⁹⁰.

Beta-actin (β -ACT) was used as a reference gene in order to normalize samples for DNA input. Additionally, the modified CpGenome™ Universal Methylated DNA was diluted in five serial dilutions (5x dilution factor) and used to generate a standard curve. For each sample, the mean quantity of *ARID1A* methylation levels was normalized to the mean quantity for the endogenous control β -ACT, according to the following formula:

$$\text{Methylation levels} = \frac{ARID1A}{\beta-ACT} \times 1000$$

Reactions were performed in 384-well plates using a Light Cycler 480 instrument (Roche Diagnostics, Mannheim, Germany). Triplicates of all samples and two negative controls were included in each plate. Briefly 2µL of modified DNA, 5 µL of SYBR Xpert Fast SYBR 2X Master Mix (GRiSP, Porto, Portugal) and 0.4µL of primers were added in each well up to a volume of 10µL. Primer sequences used and conditions are depicted in Table 2.

Table 2. Primers used in qMSP

Gene	Sequence	Annealing T/ Primer volume
<i>β-ACTIN</i>	F-5'TGGTGATGGAGGAGGTTTAGTAAGT 3' R-5'ACCAATAAAACCTACTCCTCCCTTAA 3'	60/ 0.4μL
<i>ARID1A</i>	F- 5' CGGCGTAGGTTTTAGAGATGC 3' R- 5' ACGAAACGAACGCAAACCG 3'	60/ 0.4μL

IMMUNOHISTOCHEMISTRY

Briefly, tumour blocks were sectioned at a thickness of 4 μm, deparaffinised in xylene and hydrated through a graded alcohol series. Antigen retrieval was achieved by microwave or water bath during 20 minutes in ethylenediamine tetra acetic acid (EDTA) buffer. Endogenous peroxidases were inactivated by 0.6% hydrogen peroxide (H₂O₂) and blocking of antibody nonspecific binding was achieved through incubation with horse serum (Vector Laboratories S-2000 Normal Horse Serum, concentrated; 20 mL) in a 1:50 dilution, for 20 minutes each. Slides were then incubated, according to optimized conditions, with the primary antibody (Table 3).

Table 3. Primary antibodies used in IHC and conditions.

Antibody	Antigen retrieval method	Buffer	Dilution	Incubation time	DAB (min)
ARID1A (sc-32761) Santa Cruz Biotechnology	Microwave	EDTA (pH=8)	1:800	Overnight at 4°C	10
ANXA1 (713400) Invitrogen	Water bath	EDTA (pH=8)	1:1500	1h at room temperature	7

The slides were incubated with post primary block and then with polymer (Novocastra Novolink™) for 30 minutes each. Following, diaminobenzidine tetrahydrochloride (DAB), diluted in phosphate-buffered saline (PBS), was used as a chromogen. Lastly, the slides were counterstained with haematoxylin and mounted with Entellan® (Merck-Millipore, Germany).

Normal oesophagus tissue was used as an external positive control for ANXA1 and normal cervix for ARID1A antibody.

IMMUNOHISTOCHEMISTRY SCORING

In each case, lymphocytes were used as an internal positive control, for the evaluation of both antibodies.

For ANXA1 immunostaining, percentage of cells stained and cytoplasmatic intensity of staining were assessed. Intensity was scored from 0 to 3+, representing negative to strong staining. A score of 3+ was assigned when the intensity of staining was equivalent to that of lymphocytes. The overall score was determined as previously described⁹¹: *overall score = [(%cells with intensity score 1) x 1] + [(%cells with intensity score 2) x 2] + [(%cells with intensity score 3) x 3]*.

Concerning ARID1A, since all samples had approximately 90-100% of cells stained, only intensity was evaluated. Intensity ranged from score 0 (absence) to 3+, indicative of high intensity. A score of 2+ was equivalent to lymphocytes' staining intensity.

For the statistical analysis, as no clear cut-off was defined, ANXA1 staining was categorized into "negative" and "positive" expression, considering the 75 percentile. On the other hand, ARID1A was grouped into two categories: "low intensity", comprising intensity scores 1+ and 2+, and "high intensity" comprising only the 3+ intensity score.

STATISTICAL ANALYSIS

All data were analysed using SPSS statistical software (version 24.0, Chicago, IL, USA). Non-parametric tests were used to compare ARID1A and ANXA1 immunoexpression between molecular subtypes and to evaluate associations with other clinicopathological features. Associations between proteins' immunoexpression and molecular subtypes and other clinicopathological variables were assessed by Pearson's chi-square or Fisher's exact tests.

Disease-specific (DSS), disease-free (DFS) and trastuzumab resistance-free survival (TRFS) curves were assessed through the Kaplan–Meier method. The log-rank test was computed to analyse differences in survival. Univariable Cox regression was used to assess standard clinicopathological variables and proteins' prognostic value. To understand which variables remained independent predictors of survival, a multivariable analysis was performed using the Cox proportional hazards model using the backward conditional method. A *p* value of less than 0.05 was considered statistically significant.

DFS was defined as the time between surgery date and recurrence date and DSS was defined as the time between diagnostic date and death from the disease. To perform TRFS analysis, a new variable called "resistant" was created. Women that showed radiological evidence of recurrence during Trastuzumab therapy or within 6 months after cessation, were considered resistant.

All graphs were constructed using GraphPad Prism version 6.01 for Windows (GraphPad Software, La Jolla California USA).

RESULTS

CLINICAL AND PATHOLOGICAL DATA

This study comprised 215 female patients with HER-2+ BrC, whose treatment included Trastuzumab (Table 4). Seventy eight percent of tumours were luminal B-like (HER-2+) and 22% were HER-2+ (non-luminal), as assessed by IHC assay. Most of the tumours were invasive carcinomas of no special type (NST), grade (G) 3 and stage I/II.

Table 4. Clinicopathological features of luminal-B like (HER-2+) and HER-2+ (non-luminal) BrC patients enrolled in the study.

Clinicopathological features	Luminal B-like (HER-2+)	HER-2+ (non-luminal)
Patients (n)	167	48
Age median (range)	51 (24-71)	54.5 (27-69)
Histological type (%)		
Invasive Carcinoma, no special type (NST)	128 (76.6)	43 (89.6)
Invasive Lobular Carcinoma	3 (1.8)	1 (2.1)
Other invasive carcinoma subtypes ^a	36 (21.7)	4 (8.5)
Lymphovascular invasion (%)		
No	88 (52.7)	20 (41.7)
Yes	69 (41.3)	25 (52.1)
Not determined	10 (6)	3 (6.7)
Grade (%)		
G1 & G2	84 (50.3)	13 (27.1)
G3	83 (49.7)	34 (70.8)
Not determined	-	1 (2.1)
Estrogen Receptor Status (%)		
Positive	167 (100)	-
Negative	-	48 (100)
Progesterone Receptor Status (%)		
Positive	128 (76.6)	-
Negative	39 (23.4)	48 (100)
Primary tumour (T) (%)		
T1 & T2	152 (91)	45 (93.8)
T3 & T4	13 (7.8)	3 (6.3)
Not determined	2 (1.2)	
Regional lymph node (N) (%)		
N0	67 (40.1)	18 (37.5)
N+	99 (59.3)	30 (62.5)
Not determined	1 (0.6)	-
Stage (%)		
I/II	121 (72.5)	37 (77.1)
III	45 (26.9)	11 (22.9)
Not determined	1 (0.6)	-

^a Includes medullary, mucinous and mixed type carcinoma (invasive carcinoma, NST and micropapillary carcinoma)

ARID1A PROMOTER METHYLATION STATUS

Since ARID1A's low expression might be attributable to hypermethylation, according to Zhang *et. al*⁶⁰, its promoter methylation status in our patients' cohort was assessed, however, no amplification was observed by qMSP (data not shown).

RELATIONSHIP BETWEEN ARID1A AND ANXA1

Contrarily to what was previously reported⁶⁶, no inverse association was found between ANXA1 and ARID1A immunoexpression (Figure 5B; $p=0.183$). On the contrary, tumours with ANXA1 expression seem to exhibit higher ARID1A protein levels.

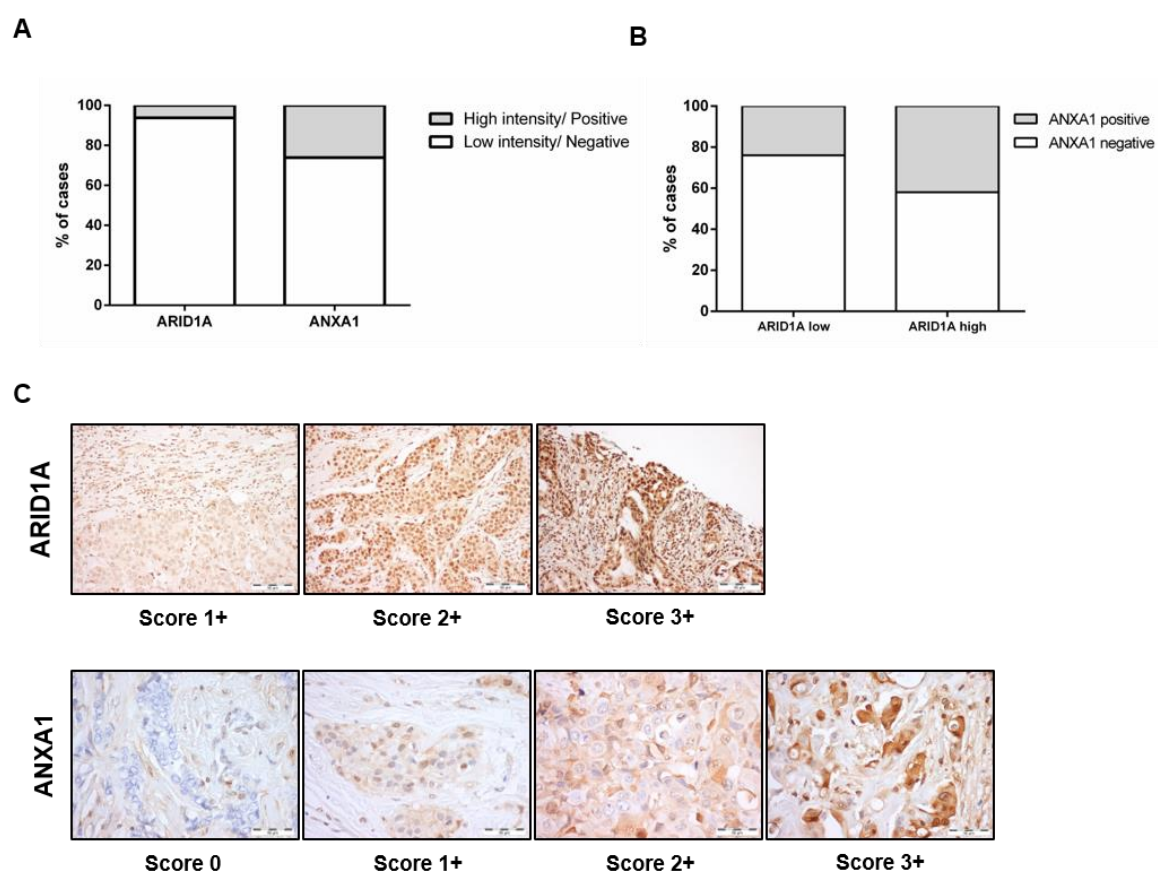


Figure 5. Relationship between ARID1A and ANXA1. Proportion of tumours with ARID1A high or low intensity and ANXA1 negativity or positivity **(A)**. Association between ARID1A and ANXA1 [Chi-square: $p=0.183$] **(B)**. Illustrative images of the different protein intensity scores **(C)**.

ARID1A AND ANXA1 EXPRESSION BY MOLECULAR SUBTYPE

Although ARID1A immunoexpression did not associate with BrC molecular subtype, HER-2+ (non-luminal) tumours depicted higher ANXA1 protein levels ($p<0.001$) than luminal B-like (HER-2+) (Figure 6).

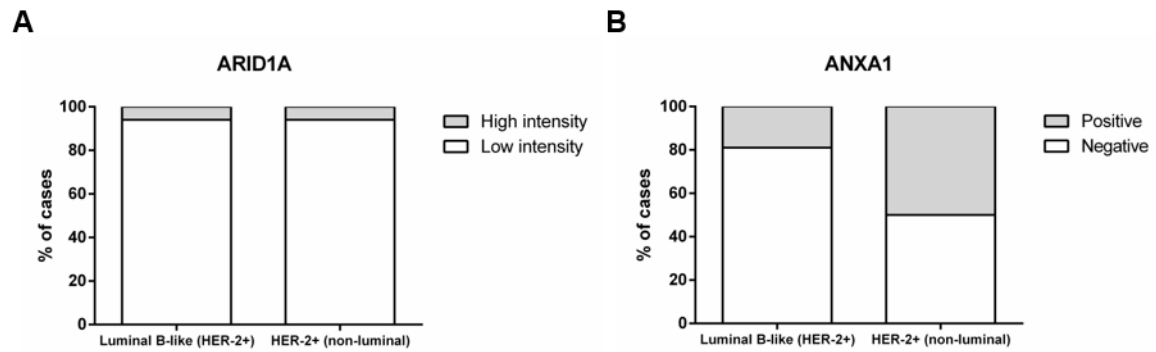


Figure 6. Distribution of ARID1A and ANXA1 immunoeexpression by molecular subtype. Percentage of cases with low and high ARID1A intensity staining score (1+ & 2+ vs 3+) [Chi-square $p=0.749$] **(A)**. Percentage of cases with or without ANXA1 expression [Chi-square $p<0.001$] **(B)**.

Additionally, no significant associations were found between both ARID1A and ANXA1 proteins' immunoeexpression and any other clinicopathological variable (age, histological type, lymphovascular invasion, grade, T stage, N stage and stage).

RESISTANCE TO TRASTUZUMAB

The value of ARID1A and ANXA1 expression as predictive biomarkers of Trastuzumab resistance was assessed using disease recurrence as a surrogate. Therefore, patients that showed radiological evidence of disease during Trastuzumab or 6 months after Trastuzumab cessation were considered resistant⁹². Since our cohort is comprised by two distinct molecular subtypes, we assessed if luminal B-like (HER-2+) and HER-2+ (non-luminal) displayed different survival, but no differences were depicted (data not shown). Furthermore, since the small number of HER-2+ (non-luminal) patients ($n=48$) would compromise the statistical power, survival analysis was performed for all the patients and not stratified by molecular subtype. Only 9 patients presented recurrent tumours within this time period and neither ARID1A nor ANXA1 expression predicted Trastuzumab resistance in this group of BrC patients (Figure 7).

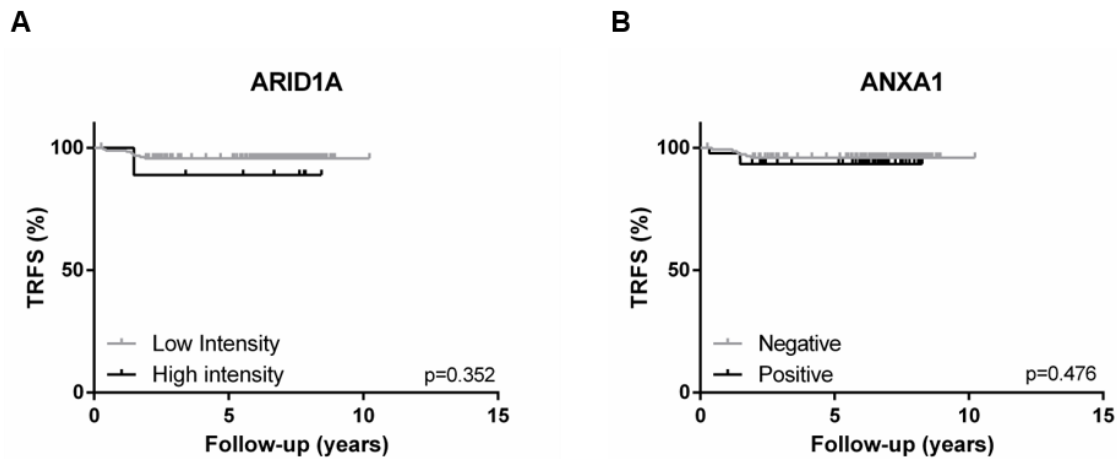


Figure 7. ARID1A and ANXA1 as predictors of Trastuzumab resistance. Trastuzumab-resistance free survival curves (Kaplan-Meier with log rank test) of ARID1A **(A)** and ANXA1 **(B)** immunoexpression, considering patients that recurred during or within 6 months after Trastuzumab cessation, as resistant.

HIGH ARID1A AND ANXA1 EXPRESSION IS ASSOCIATED WITH EARLY RECURRENCE AND SHORTER SURVIVAL

Patients' median follow-up time was 83 months. From the 215 patients included in this study, 31 (14.4%) deceased due to BrC, whereas 180 remained alive, 10 of which (4.7%) harbouring cancer.

Due to the reduced number of events and/or cases in some categories, some clinicopathologic features were grouped. Grade was grouped as (G1 & G2 vs. G3), T stage was grouped as (T1 & T2 vs. T3 & T4), N stage was grouped as (N0 vs. N1) and stage was grouped as (I & II vs. III). ANXA1 was grouped as "negative vs. positive" according with p75 final score, whereas ARID1A protein staining intensity was grouped as "1+ & 2+ vs. 3+", as previously stated.

DSS and DFS did not differ between luminal B-like (HER-2+) and HER-2+ (non-luminal) tumours (data not shown), thus survival analysis was performed for all the patients.

Lymphovascular invasion, larger tumours (T3 and T4), positive lymph nodes and stage III disease significantly associated with worse survival and recurrent disease (Supplementary Figure 1 and Supplementary Figure 2). Notably, higher expression levels of ARID1A and ANXA, not only disclosed shorter DSS, but also earlier disease recurrence (Figure 8).

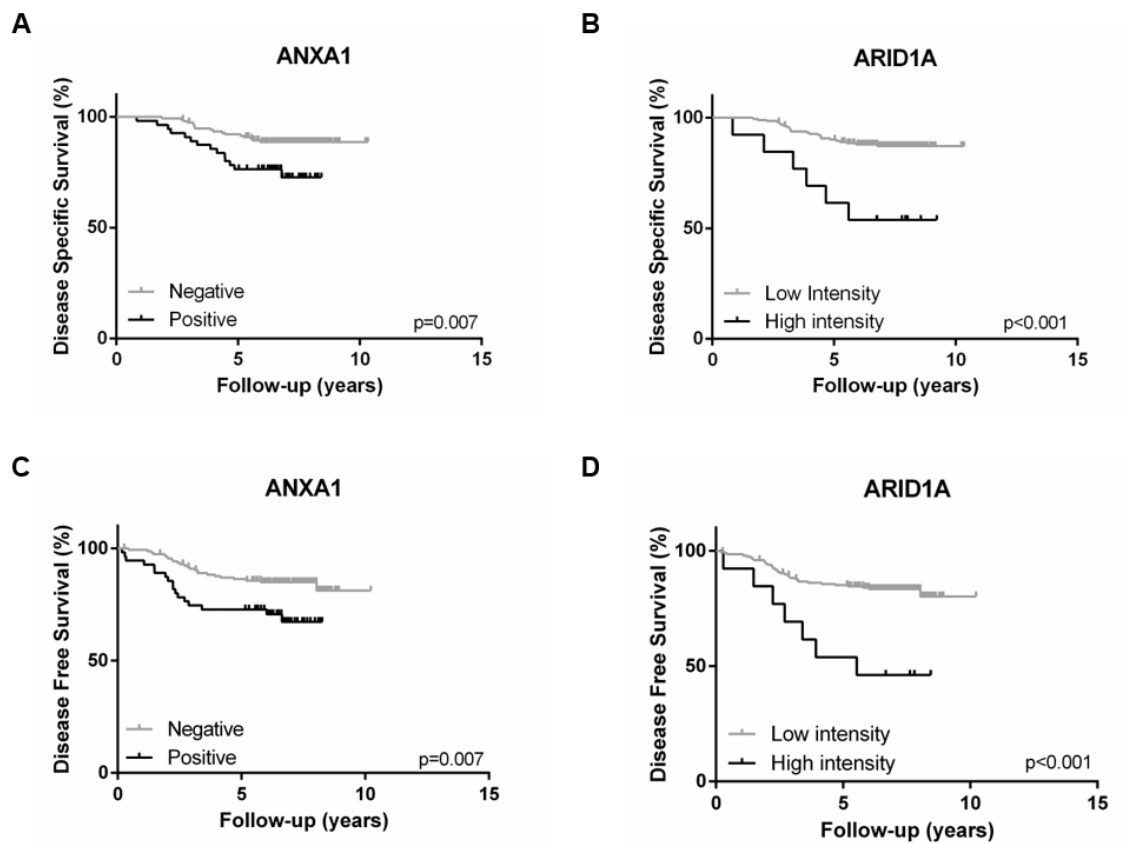


Figure 8. ARID1A and ANXA1 prognostic value. Disease-specific and disease-free survival curves (Kaplan-Meier with log rank test) of ANXA1 (A and C, respectively) and ARID1A (B and D, respectively) immunoexpression.

Cox regression analysis was also computed to assess the potential of clinicopathological variables and ANXA1 immunoexpression as survival predictors. Despite disclosing shorter DSS and DFS in univariable analysis, lymphovascular invasion, T stage, N stage and ARID1A were not included in the Cox regression analysis due to the reduced number of events in each group.

In the multivariable analysis, along with stage, ANXA1 immunoexpression independently predicted patients' DSS (Table 5). Strikingly, patients with ANXA1 positive tumours have, approximately, 3 times more probability of dying from BrC than those without expression.

Table 5. Cox regression models assessing the potential of clinical variables and ANXA1 immunoexpression in the prediction of disease-specific survival.

Disease -specific survival	Variable	Hazard Ratio (HR)	95% CI for HR	<i>p value</i>
Univariable	Stage I&II	1	2.415-10.268	<0.001
	III	4.980		
	ANXA1 Negative	1	1.259-5.189	0.009
	Positive	2.557		
Multivariable	Stage I&II	1	2.374-10.093	<0.001
	III	4.895		
	ANXA1 Negative	1	1.309-5.393	0.007
	Positive	2.658		

Additionally, ANXA1 positivity independently predicted shorter time to recurrence (Table 6).

Table 6. Cox regression models assessing the potential of clinical variables and ANXA1 immunoexpression in the prediction of disease-free survival.

Disease-free survival	Variable	Hazard Ratio (HR)	95% CI for HR	<i>p value</i>
Univariable	Stage I&II	1	2.215-7.490	<0.001
	III	4.073		
	ANXA1 Negative	1	1.239-4.304	0.008
	Positive	2.309		
Multivariable	Stage I&II	1	2.270-7.817	<0.001
	III	4.213		
	ANXA1 Negative	1	1.296-4.499	0.005
	Positive	2.415		

Overall, our results indicate that ANXA1 might contribute to adverse outcomes related to survival and recurrence in HER-2+ BrC.

DISCUSSION

Regardless of the great efforts made for improving BrC patients' management, it still remains the deadliest cancer among women¹. Genomic and expression profiling analysis granted an insight of the true molecular features of a tumour, improving the understanding of its behaviour⁹³. ESMO distinguishes four intrinsic subtypes (luminal A, luminal B, HER-2 enriched and "basal-like") that display different patterns of gene expression, as well as different prognosis and outcomes¹⁹. Additionally, the recognition of these entities also aids clinicians in the choice of treatment to assign to each patient⁹⁴. Indeed, IHC for ER, PR and HER-2 provide a surrogate intrinsic classification of tumours identifying, among others, luminal B-like (HER-2+) and HER-2+ (non-luminal) tumours, which are characterized by an overexpression of the HER-2 receptor¹⁹. The standard of care of these HER-2+ subtype patients includes the use of Trastuzumab^{19,95}. However, recurrence and disease progression dampen the effectiveness granted by this therapy. Hence, new biomarkers that might better identify the BrC patients who are most likely to benefit from Trastuzumab are urgently needed⁶¹.

ARID1A and ANXA1 expression were suggested to associate with Trastuzumab resistance. Specifically, ANXA1 was also implicated in signalling pathways that impact Trastuzumab effectiveness^{64,65,96}.

In parallel, previous studies associated ARID1A loss with worse prognosis in several cancers^{97,84,98}, although its function in BrC is not entirely understood^{99,100,87}. Furthermore, those studies have only assessed a limited number of HER-2+ samples^{87,101}. Indeed, most statistically significant associations of ARID1A and outcome were established for TNBC BrC¹⁰².

Concerning ANXA1, its expression was associated with BrC aggressiveness, progression, higher metastatic potential and also with triple-negative phenotype^{65,103,104}. Remarkably, ANXA1 was reported to modulate cell adhesion and motility through transforming growth factor- β (TGF β) activation, thus, leading to EMT switch, supporting the earlier described BrC features^{105,59,106,107}. Along with this, TGF β also activates PI3K signalling pathway, a mechanism implicated in tumour cells' unresponsiveness to Trastuzumab. Specifically, AKT activation was implicated in such resistance, being associated with worse prognosis in some types of cancer, including BrC^{108,109,110,111}.

In this study we analysed 215 HER-2+ BrC patients' specimens to investigate the value of ARID1A and ANXA1 expression on clinical outcome and prediction of Trastuzumab resistance.

It was previously reported by Zhang *et al.*⁸⁸ that ARID1A mRNA loss could be attributed to promoter hypermethylation. However, in our cohort, no aberrant methylation was found for the same promoter region studied before. Nonetheless, to our knowledge, that is the only study reporting that ARID1A downregulation associated with promoter

methylation in BrC. The discrepancies observed might result from the different methodologies used. In our study qMSP was performed instead of methylated DNA immunoprecipitation followed by PCR. Fresh frozen tissue samples, used by the other research team, are considered ideal for molecular analysis¹¹², however, FFPE tissues were used in our study. FFPE samples are a difficult material to work with in molecular analysis since the fixation process leads to DNA fragmentation which can affect the quality of downstream applications¹¹³. Despite being a gold standard procedure to study DNA methylation, bisulphite conversion, also performed in our study, is a harsh chemical reaction that further degrades DNA¹¹⁴. Furthermore, the small number of samples included in the study (n=38) by the other research team may have also introduced a bias in the analysis. Hence, future studies highlighting this topic must be performed.

Importantly, HER-2+ (non-luminal) tumours depicted higher ANXA1 expression, which is consistent with previous studies reporting a correlation between ANXA1 and HR negative BrC subtypes^{105,115}. This finding suggests that the oncogenic role of ANXA1 in BrC may be attenuated in tumours expressing HR and that its expression may be more relevant in HER-2+ (non-luminal) tumours. Nevertheless, the relatively small size of our cohort implies that further validation in a larger patient cohort is required.

Contrarily to what was reported by Berns *et al.*⁶⁶, an inverse association between ARID1A and ANXA1 was not depicted in our cohort. The authors primarily based their findings in results obtained through functional assays in cell lines. Additionally, they resorted to a TCGA panel of BrC patients and found the same association between ARID1A and ANXA1 protein expression. However, this protein analysis differs from ours since they used data from RPPA instead of IHC and the series comprised all subtypes of BrC patients, and not only HER-2+. Furthermore, given that ARID1A “high intensity” category, in our study, only comprises 13 cases, this may be accountable for the observed results.

No significant differences were found concerning these proteins’ value as predictors of Trastuzumab resistance. Previous studies addressing this subject have used recurrence-free survival as a surrogate definition of resistance to Trastuzumab⁶⁶. In our opinion, this is not the most accurate definition of resistance since patients may experience recurrence many years after receiving Trastuzumab. Given that Trastuzumab is usually administered for a relatively short period of time (1 year), we consider that recurrences during this period or during a brief period (for instance, 6 months) after cessation should give a more accurate meaning of unresponsiveness to Trastuzumab. Indeed, this was the time frame considered for patient inclusion in a clinical trial evaluating the value of another treatment for HER-2+ patients that either recurred or progressed on Trastuzumab⁹². Hence, a proper definition of “Trastuzumab resistant patients” must be created in order to standardize future studies regarding prediction biomarkers evaluation.

Our data revealed that patients with higher ARID1A and ANXA1 expressing tumours showed increased recurrence risk and higher risk of dying from this disease. Hence, these proteins' expression may be useful as recurrence and survival biomarkers. However, it should be considered that, due to the small number of events, ARID1A was not included in the multivariable analysis, thus, further studies with larger cohorts must address this issue. Moreover, since ARID1A and ANXA1 contribute to poorer prognosis and pre-exist in treatment-naïve tumours, they may be important to identify which HER-2+ patients display high expression and tailor a different therapeutic approach for them. Since it is proposed that ANXA1 render Trastuzumab resistance through activation of AKT, additional therapeutics targeting Pi3K/AKT/mTOR pathway should be considered for this subset of patients. Indeed, other studies attempted to associate AKT activation and response to Trastuzumab. A recent study evaluating the effect of carboxyl-terminal modulator protein (CTMP) in Trastuzumab resistance, showed that AKT activation is implicated in tumours' unresponsiveness, in HER-2+ BrC patients¹¹⁰. Moreover, the authors also found that higher levels of CTMP were related with worse survival in HER-2 enriched patients. In the same line, by functional assays, higher phosphorylated AKT levels were associated with resistance to Trastuzumab. Hence, AKT signalling or its downstream effectors inhibition may also be used as a therapeutic approach to overcome Trastuzumab resistance. Also, notwithstanding these proteins' role in BrC initiation and progression being far from understood, ANXA1's ability to discriminate luminal B-like (HER-2+) and HER-2+ (non-luminal) subtypes, contributes to a better patient stratification regarding treatment strategy. Despite many studies associate ARID1A loss with worse prognosis, whereas our results suggest otherwise, few of them were performed in BrC tissue, especially in HER-2+ subtypes. Also, our results are in line with one study that reported high ARID1A expression in invasive and mucinous carcinomas in TCGA dataset, thereby suggesting its involvement in breast carcinogenesis¹¹⁶. Therefore, this protein role in BrC is not yet established.

The main limitations of this study were the relatively small number of HER-2+ (non-luminal) tumours and the limited number of recurrences and deaths observed. It must also be stressed that only 13 samples comprised ARID1A's "high expression group" and, for that reason, further studies with a larger number of samples should be performed to increase the analysis' statistical power. Also, it should be recalled, once again, that current guidelines that specifically define resistance to Trastuzumab are still lacking. To overcome that restraint, we have used the definition reported by the EMILIA clinical trial, which focused on the best treatment to be assigned to HER-2+ locally advanced or metastatic BrC patients who stopped responding to Trastuzumab⁹².

To our knowledge, this is the first study assessing the prognostic and prediction value of ARID1A and ANXA1 proteins' in HER-2+ BrC patients whose treatment included Trastuzumab.

Nonetheless, large- scale, multicentric and extended follow-up studies are required to better understand the prognostic and predictive value of these proteins in identifying BrC patients that will develop resistance to Trastuzumab. The need to identify and validate molecular predictors of response to Trastuzumab that might allow to better stratify HER-2+ patients is a major quest, since new drugs that may be able to overcome resistance already exist^{45,92}.

CONCLUSIONS AND FUTURE PERSPECTIVES

The need to identify and validate molecular predictors of response to trastuzumab that might allow a better stratification of HER-2+ patients is urgently required, since new drugs that may be able to overcome resistance already exist^{45,117}. The identification of the patients that are most likely to recur will improve life quality and, additionally, given the high cost of Trastuzumab, it will also ease the economic burden on healthcare systems.

In this study we were able to confirm the value of ANXA1 as a marker of worse prognosis in BrC. Indeed, patients whose tumours show ANXA1 protein expression are approximately 3 times more likely to recur and die of BrC. Furthermore, we addressed this question in a cohort composed of HER-2+ patients treated with a Trastuzumab-based therapy, which required elucidation. Large-scale multicentric prospective studies are demanded to validate ANXA1 and, specially, ARID1A value in HER-2+ BrC outcome. Since ANXA1 showed to be a promising prognostic biomarker, it would also be interesting to assess its mRNA levels in fresh-frozen tissue, which provides more reliable results in molecular analysis than FFPE tissue. A major goal would be to create a prognostic test based on this gene's expression; however, rigorous validation is mandatory.

Overall, our results support a prognostic value of ANXA1 in HER-2+ BrC patients treated with a Trastuzumab-based therapy. If standardization and validation is achieved, ANXA1's assessment will provide a useful clinical asset for patient stratification and prognosis.

REFERENCES

1. Bray, F., Ferlay, J. & Soerjomataram, I. (2018) Global Cancer Statistics 2018 : GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA. Cancer J. Clin.* **00**, 1–31.
2. E. Senkus, et al. (2015) Clinical practice guidelines Primary breast cancer : ESMO Clinical Practice Guidelines for diagnosis , treatment and follow-up clinical practice guidelines. *Ann. Oncol.* **26**, 8–30.
3. DeSantis, C. E. et al. (2015) International Variation in Female Breast Cancer Incidence and Mortality Rates. *Cancer Epidemiol. Biomarkers Prev.* **24**, 1495 LP – 1506.
4. Tao, Z. Q. et al. (2015) Breast Cancer: Epidemiology and Etiology. *Cell Biochem. Biophys.* **72**, 333–338.
5. Becker, S. (2015) A historic and scientific review of breast cancer : The next global healthcare challenge. *Int. J. Gynecol. Obstet.* **131**, S36–S39.
6. Mcpherson, K., Steel, C. M. & Dixon, J. M. (2000) Breast cancer — epidemiology , risk factors , and genetics Risk factors for breast cancer. *BMJ* **321**:624,
7. Dumitrescu, R. G. & Cotarla, I. (2005) Understanding breast cancer risk - where do we stand in 2005 ? *J. Cell. Mol. Med.* **9**, 208–221.
8. Powers, S., Zhu, W., Wu, S. & Hannun, Y. A. (2015) Substantial contribution of extrinsic risk factors to cancer development. *Nature* **529**, 43.
9. Hulka, B. S. et al. (2001) Breast cancer : hormones and other risk factors. *Maturitas* **38**, 103–113.
10. Vogel, V. G. (2012) *Epidemiology of Breast Cancer. Breast Pathology* (Elsevier, 2012).
11. Fackenthal, J. D. & Olopade, O. I. (2007) Breast cancer risk associated with BRCA1 and BRCA2 in diverse populations. *Nat. Rev. Cancer* **7**,
12. Schon, K. & Tischkowitz, M. (2018) Clinical implications of germline mutations in breast cancer : *Breast Cancer Res. Treat.* **167**, 417–423.
13. Zhang, H. Y., Liang, F., Jia, Z. H. I. L., Song, S. & Jiang, Z. (2013) PTEN mutation , methylation and expression in breast cancer patients. *Oncol. Lett.* **6**, 161–168.
14. Oh, H. et al. (2017) Breast cancer risk factors in relation to estrogen receptor, progesterone receptor, insulin-like growth factor-1 receptor, and Ki67 expression in normal breast tissue. *NPJ Breast Cancer* **3**, 1–7.
15. Kaaks, R. et al. (2005) Postmenopausal serum androgens , oestrogens and breast cancer risk : the European prospective investigation into cancer and nutrition. *Endocr. Relat. Cancer* **12**, 1071–1082.
16. Russo, J., Moral, R., Balogh, G. A., Mailo, D. & Russo, I. H. (2005) The protective role of pregnancy in breast cancer. *Breast Cancer Res.* **7**, 131–142.

17. Donaldson, M. S. (2004) Nutrition and cancer : A review of the evidence for an anti-cancer diet. *Nutr. J.* **21**, 1–21.
18. Brown, J. C., Winters-stone, K., Lee, A. & Schmitz, K. H. (2012) Cancer , Physical Activity , and Exercise. *Compr. Physiol.* **2**, 2775–2809.
19. Cardoso, F. *et al.* (2019) Early breast cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann. Oncol.* **30**, 1194–1220.
20. Marmot, M. G. *et al.* (2012) The benefits and harms of breast cancer screening: An independent review. *Lancet* **380**, 1778–1786.
21. Tavassoli FA., D. P. (2003) Pathology and Genetic of Tumours of the Breast and Female Genital Organs, WHO Classification of Tumours. *Pathol. Genet.* **4**,
22. Makki, J. (2015) Diversity of Breast Carcinoma : Histological Subtypes and Clinical Relevance. *Clin. Med. insights Pathol.* **8**, 23–31.
23. Govindan R, D. V. (2009) DeVita, Hellman, and Rosenberg’s Cancer: Principles & Practice of Oncology Review. *Lippincott Williams & Wilkins*
24. Lal, S., Reed, A. E. M., Luca, X. M. De & Simpson, P. T. (2017) Molecular signatures in breast cancer A . Tumour Grade N : Lymph Node Involvement M : Metastasis. *Methods* **131**, 135–146.
25. Giuliano, A. E. *et al.* (2017) Breast Cancer-Major changes in the American Joint Committee on Cancer eighth edition cancer staging manual. *CA. Cancer J. Clin.* **67**, 290–303.
26. Amin MB, et al. (2017) *AJCC Cancer Staging Manual*. (Springer International Publishing, 2017).
27. Yip, C. & Rhodes, A. (2014) Estrogen and progesterone receptors in breast cancer. *Futur. Med.* **10**, 2293–2301.
28. Dai, X., Xiang, L., Li, T. & Bai, Z. (2016) Cancer Hallmarks , Biomarkers and Breast Cancer Molecular Subtypes. *J. Cancer* **7**, 1281–1294.
29. Gutierrez, M. C. *et al.* (2018) Molecular Changes in Tamoxifen-Resistant Breast Cancer : Relationship Between Estrogen Receptor , HER-2 , and p38 Mitogen-Activated Protein Kinase. **23**, 2469–2476.
30. Arteaga, C. L. *et al.* (2011) Treatment of HER2-positive breast cancer : current status and future perspectives. *Nat. Rev. Clin. Oncol.* **9**, 16–32.
31. Carolina Gutierrez, R. S. (2011) HER 2: Biology, Detection, and Clinical Implications. *Arch. Pathol.* **135**, 55–62.
32. Kathleen I. Pritchard, et al. (2013) HER2 and Responsiveness of Breast Cancer to Adjuvant Chemotherapy. *N. Engl. J. Med.* **354**, 485–493.
33. Dowsett, M. *et al.* (2011) Assessment of Ki67 in Breast Cancer: Recommendations from the international Ki67 in breast cancer working Group. *J. Natl. Cancer Inst.* **103**,

1656–1664.

34. Penault-Llorca, F. & Radosevic-Robin, N. (2017) Ki67 assessment in breast cancer: an update. *Pathology* **49**, 166–171.
35. Gyorffy, B. *et al.* (2015) Multigene prognostic tests in breast cancer: Past, present, future. *Breast Cancer Res.* **17**, 1–7.
36. Verma, A., Kaur, J. & Mehta, K. (2015) Molecular oncology update: Breast cancer gene expression profiling. *Asian J. Oncol.* **1**, 65–72.
37. Sørlie, T. (2004) Molecular portraits of breast cancer: Tumour subtypes as distinct disease entities. *Eur. J. Cancer* **40**, 2667–2675.
38. Prat, A. *et al.* (2015) Clinical implications of the intrinsic molecular subtypes of breast cancer. *The Breast* **24**, 26–35.
39. Perou, C. M. *et al.* (2000) Molecular portraits of human breast tumours. *Nature* **533**, 747–752.
40. Ozlem Yersal. (2014) World Journal of Clinical Oncology © 2014. *World J. Clin. Oncol.* **5**, 412–425.
41. Kittaneh, M. & Montero, A. J. (2013) Biomarkers in Cancer Molecular Profiling for Breast Cancer : A Comprehensive Review. *Biomark. Cancer* **5**, 61–70.
42. Dai, X. *et al.* (2015) Breast cancer intrinsic subtype classification, clinical use and future trends. *Am J Cancer Res* **5**, 2929–2943.
43. Cooke, T., Reeves, J., Lanigan, a & Stanton, P. (2001) HER2 as a prognostic and predictive marker for breast cancer. *Ann. Oncol.* **12 Suppl 1**, S23–S28.
44. Natalija Dedi Plaveti *et al.* (2012) Role of HER2 signaling pathway in breast cancer : biology , detection and therapeutical implications. *Period. Biol.* **114**, 505–510.
45. Lavaud, P. & Andre, F. (2014) Strategies to overcome trastuzumab resistance in HER2-overexpressing breast cancers : focus on new data from clinical trials. *BMC Med.* **12**, 1–10.
46. Vu, T. & Claret, F. X. (2012) Trastuzumab: Updated Mechanisms of Action and Resistance in Breast Cancer. *Front. Oncol.* **2**, 1–6.
47. Ross JS1, F. J. (1998) The HER-2/neu Oncogene in Breast Cancer: Prognostic Factor, Predictive Factor, and Target for Therapy. *Stem Cells* **16**, 413–428.
48. Eroles, P., Bosch, A., Pérez-fidalgo, J. A. & Lluch, A. (2012) Molecular biology in breast cancer : Intrinsic subtypes and signaling pathways. *Cancer Treat. Rev.* **38**, 698–707.
49. Nounou, M. I. *et al.* (2015) Breast Cancer: Conventional Diagnosis and Treatment Modalities and Recent Patents and Technologies. *Breast Cancer (Auckl)*. **9**, 17–34.
50. Kokko, R. *et al.* (2006) Adjuvant Docetaxel or Vinorelbine with or without Trastuzumab for Breast Cancer. *N. Engl. J. Med.* **354**, 809–820.

51. O'Brien, N. A. *et al.* (2010) Activated Phosphoinositide 3-Kinase/AKT Signaling Confers Resistance to Trastuzumab but not Lapatinib. *Mol. Cancer Ther.* **9**, 1489–1502.
52. Arnould, L. *et al.* (2006) Trastuzumab-based treatment of HER2-positive breast cancer: An antibody-dependent cellular cytotoxicity mechanism? *Br. J. Cancer* **94**, 259–267.
53. Nahta, R. & Esteva, F. J. (2007) Trastuzumab: Triumphs and tribulations. *Oncogene* **26**, 3637–3643.
54. Vogel, C. L. *et al.* (2003) Efficacy and Safety of Trastuzumab as a Single Agent in First-Line Treatment of HER2-Overexpressing Metastatic Breast Cancer. *J. Clin. Oncol.* **20**, 719–726.
55. Et.al, M. J. P.-G. (2005) Trastuzumab after Adjuvant Chemotherapy in HER2-positive Breast Cancer. *N. Engl. J. Med.* **353**, 1659–1672.
56. Bryant, J. *et al.* (2005) Trastuzumab plus Adjuvant Chemotherapy for Operable HER2-Positive Breast Cancer. *N. Engl. J. Med.* **353**, 1673–1684.
57. Nahta, R. & Esteva, F. J. (2006) Molecular mechanisms of trastuzumab resistance. *Breast Cancer Res.* **8**, 1–8.
58. Carraway, K. & et. al. (2001) Muc4/sialomucin complex in the mammary gland and breast cancer. *J. Mammary Gland Biol. Neoplasia* **6**, 323–337.
59. Nahta, R., Yu, D., Hung, M. C., Hortobagyi, G. N. & Esteva, F. J. (2006) Mechanisms of disease: Understanding resistance to HER2-targeted therapy in human breast cancer. *Nat. Clin. Pract. Oncol.* **3**, 269–280.
60. Yuhong, L. & et. al. (2001) Insulin-Like Growth Factor-I Receptor Signaling and Resistance to Trastuzumab (Herceptin). *J. Natl. Cancer Inst.* **93**, 1852–1857.
61. Nagata, Y. *et al.* (2004) PTEN activation contributes to tumor inhibition by trastuzumab, and loss of PTEN predicts trastuzumab resistance in patients. *Cancer Cell* **6**, 117–127.
62. Yakes, F. M. *et al.* (2002) Herceptin-induced inhibition of phosphatidylinositol-3 kinase and Akt is required for antibody-mediated effects on p27, cyclin D1, and antitumor action. *Cancer Res.* **62**, 4132–4141.
63. Berns, K. *et al.* (2007) A Functional Genetic Approach Identifies the PI3K Pathway as a Major Determinant of Trastuzumab Resistance in Breast Cancer. *Cancer Cell* **12**, 395–402.
64. Sonnenblick, A. *et al.* (2015) Integrative proteomic and gene expression analysis identify potential biomarkers for adjuvant trastuzumab resistance : analysis from the Fin-her phase III randomized trial. *Oncotarget* **6**, 30306–30316.
65. de Graauw, M. *et al.* (2010) Annexin A1 regulates TGF- signaling and promotes

- metastasis formation of basal-like breast cancer cells. *Proc. Natl. Acad. Sci.* **107**, 6340–6345.
66. Berns, K. *et al.* (2016) Loss of ARID1A activates ANXA1, which serves as a predictive biomarker for trastuzumab resistance. *Clin. Cancer Res.* **22**, 5238–5248.
 67. Nebbioso, A., Tambaro, F. P., Dell'Aversana, C. & Altucci, L. (2018) Cancer epigenetics: Moving forward. *PLoS Genet.* **14**, 1–25.
 68. Dupont, C., Armant, D. R. & Brenner, C. A. (2009) Epigenetics: Definition, mechanisms and clinical perspective. *Semin. Reprod. Med.* **27**, 351–357.
 69. Tsai, H. C. & Baylin, S. B. (2011) Cancer epigenetics: Linking basic biology to clinical medicine. *Cell Res.* **21**, 502–517.
 70. Robertson, K. D. (2005) DNA methylation and human disease. *Nat. Rev. Genet.* **6**, 597–610.
 71. Flavahan, W. A., Gaskell, E. & Bernstein, B. E. (2017) Epigenetic plasticity and the hallmarks of cancer. *Science (80-.).* **357**, 266–274.
 72. Mukesh Verma. (2015) *Cancer Epigenetics - Risk Assessment, Diagnosis, Treatment, and Prognosis*. (Humana Press, 2015).
 73. Sandoval, J. & Esteller, M. (2012) Cancer epigenomics: Beyond genomics. *Curr. Opin. Genet. Dev.* **22**, 50–55.
 74. Voyias P, Patel A, A. R. (2016) Epigenetic Biomarkers of Disease. in *General Medical Aspects of Epigenetics* 159–176. (Elsevier, 2016).
 75. Marta Kulis, M. E. (2010) DNA Methylation and Cancer. *Adv. Genet.* **70**, 27–56.
 76. Ito S, *et al.* (2011) Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Pharmacol. Res. Commun.* **2**, 1300–1303.
 77. Cervera, R., Ramos, A., Lluch, A. & Climent, J. (2015) *DNA Methylation in Breast Cancer. Epigenetic Biomarkers and Diagnostics* (Elsevier Inc., 2015). doi:10.1016/B978-0-12-801899-6.00015-2
 78. Costa-Pinheiro P, Montezuma D, Henrique R, J. C. (2015) Diagnostic and prognostic epigenetic biomarkers in cancer. *Epigenomics* **7**, 1003–1015.
 79. Dworkin, A. M., Huang, T. H. M. & Toland, A. E. (2009) Epigenetic alterations in the breast: Implications for breast cancer detection, prognosis and treatment. *Semin. Cancer Biol.* **19**, 165–171.
 80. Saha, A., Wittmeyer, J. & Cairns, B. R. (2006) Chromatin remodelling: the industrial revolution of DNA around histones. *Nat. Rev. Mol. Cell Biol.* **7**, 437–447.
 81. Wilson, B. G. & Roberts, C. W. M. (2011) SWI / SNF nucleosome remodellers and cancer. *Nat. Rev. Cancer* **11**, 481–492.
 82. Wu, J. N. & Roberts, C. W. M. (2013) ARID1A Mutations in Cancer: Another Epigenetic Tumor Suppressor? *Cancer Discov.* **3**, 35–44.

83. Wiegand, K. C., Shah, S. P., Al-Agha, O. M., Zhao, Y. & Al, E. (2011) ARID1A Mutations in Endometriosis-Associated Ovarian Carcinomas. *N Engl J Med* **363**, 1532–1543.
84. Cho, H. *et al.* (2013) Loss of ARID1A/BAF250a expression is linked to tumor progression and adverse prognosis in cervical cancer. *Hum. Pathol.* **44**, 1365–1374.
85. Bin, G. G. & Lien Mao, T. (2011) Mutation and loss of expression of ARID1A in uterine low-grade endometrioid carcinoma. *Am. J. Surg. Pathol.* **35**, 625–632.
86. Wang, K. *et al.* (2011) Exome sequencing identifies frequent mutation of ARID1A in molecular subtypes of gastric cancer. *Nat. Genet.* **43**, 1219–1223.
87. Mamo, A. *et al.* (2011) An integrated genomic approach identifies ARID1A as a candidate tumor-suppressor gene in breast cancer. *Oncogene* **31**, 2090–2100.
88. Zhang, X. *et al.* (2013) Promoter Hypermethylation of ARID1A Gene Is Responsible for Its Low mRNA Expression in Many Invasive Breast Cancers. *PLoS One* **8**, 1–6.
89. SR, L., IO, E., Schnitt, S. & Al., E. (2012) *WHO Classification of Tumours*. (2012).
90. Zhang, X. *et al.* (2013) Promoter Hypermethylation of ARID1A Gene Is Responsible for Its Low mRNA Expression in Many Invasive Breast Cancers. *PLoS One* **8**,
91. Yom, C. K. *et al.* (2011) Clinical significance of annexin A1 expression in breast cancer. *J. Breast Cancer* **14**, 262–268.
92. Verma, S. *et al.* (2012) Trastuzumab emtansine for HER2-positive advanced breast cancer. *N. Engl. J. Med.* **367**, 1783–91.
93. Sørlie, T. *et al.* (2001) Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 10869–10874.
94. Eliyatkin, N., Yalcin, E., Zengel, B., Aktaş, S. & Vardar, E. (2015) Molecular Classification of Breast Carcinoma: From Traditional, Old-Fashioned Way to A New Age, and A New Way. *J. Breast Heal.* **11**, 59–66.
95. Loibl, S. & Gianni, L. (2016) HER2-positive breast cancer. *Lancet* **6736**, 1–15.
96. Sun, M. *et al.* (2013) Potential role of ANXA1 in cancer. *Futur. Oncol.* **9**, 1773–1793.
97. Bosse, T. *et al.* (2013) Loss of ARID1A expression and its relationship with PI3K-Akt pathway alterations, TP53 and microsatellite instability in endometrial cancer. *Mod. Pathol.* **26**, 1525–1535.
98. Ang, X. W. *et al.* (2004) Expression of p270 (ARID1A), a component of human SWI/SNF complexes in human tumors. *Int. J. Cancer* **112**, 636–642.
99. Zhao, J., Liu, C. & Zhao, Z. (2014) ARID1A: A potential prognostic factor for breast cancer. *Tumor Biol.* **35**, 4813–4819.
100. Cho, H. D. *et al.* (2016) Loss of Tumor Suppressor ARID1A Protein Expression Correlates with Poor Prognosis in Patients with Primary Breast Cancer. *J. Breast*

Cancer **18**, 339–346.

101. International, T. *et al.* (2012) Frequent low expression of chromatin remodeling gene ARID1A in breast cancer and its clinical significance. *Cancer Epidemiol.* **36**, 288–293.
102. Lin, Y.-F. *et al.* (2018) High-level expression of ARID1A predicts a favourable outcome in triple-negative breast cancer patients receiving paclitaxel-based chemotherapy. *J. Cell. Mol. Med.* **22**, 2458–2468.
103. Okano, M. *et al.* (2015) Upregulated Annexin A1 promotes cellular invasion in triple-negative breast cancer. *Oncol. Rep.* **33**, 1064–1070.
104. Zóia, M. A. P. *et al.* (2019) Inhibition of triple-negative breast cancer cell aggressiveness by cathepsin D blockage: Role of annexin A1. *Int. J. Mol. Sci.* **20**,
105. Sobral-Leite, M. *et al.* (2015) Annexin A1 expression in a pooled breast cancer series: Association with tumor subtypes and prognosis. *BMC Med.* **13**, 1–11.
106. Porta, C., Paglino, C. & Mosca, A. (2014) Targeting PI3K / Akt / mTOR signaling in cancer. *Front. Oncol.* **4**, 1–11.
107. Nahta, R., Shabaya, S., Ozbay, T. & Rowe, D. (2012) Personalizing HER2-Targeted Therapy in Metastatic Breast Cancer Beyond HER2 Status: What We Have Learned from Clinical Specimens. *Curr. Pharmacogenomics Person. Med.* **7**, 263–274.
108. Tang, L. *et al.* (2018) NES1/KLK10 promotes trastuzumab resistance via activation of PI3K/AKT signaling pathway in gastric cancer. *J. Cell. Biochem.* **119**, 6398–6407.
109. Jin, M. H. *et al.* (2017) Resistance Mechanism against Trastuzumab in HER2-Positive Cancer Cells and Its Negation by Src Inhibition. *Mol. Cancer Ther.* **16**, 1145–1154.
110. Chen, Y.-C. *et al.* (2017) CTMP, a predictive biomarker for trastuzumab resistance in HER2-enriched breast cancer patient. *Oncotarget* **8**, 29699–29710.
111. Clark, A. S., West, K., Streicher, S. & Dennis, P. A. (2002) Constitutive and inducible Akt activity promotes resistance to chemotherapy, trastuzumab, or tamoxifen in breast cancer cells. *Mol. Cancer Ther.* **1**, 707–717.
112. Ripoli, F. L. *et al.* (2016) A comparison of fresh frozen vs. Formalin-fixed, paraffin-embedded specimens of canine mammary tumors via branched-DNA assay. *Int. J. Mol. Sci.* **17**, 1–11.
113. Ludgate, J. L. *et al.* (2017) A streamlined method for analysing genome-wide DNA methylation patterns from low amounts of FFPE DNA. *BMC Med. Genomics* **10**, 1–10.
114. Tanaka, K. & Okamoto, A. (2007) Degradation of DNA by bisulfite treatment. *Bioorganic Med. Chem. Lett.* **17**, 1912–1915.
115. Cao, Y. *et al.* (2008) Loss of annexin A1 expression in breast cancer progression.

Appl. Immunohistochem. Mol. Morphol. **16**, 530–534.

116. Wu, Y., Gu, Y., Guo, S., Dai, Q. & Zhang, W. (2016) Expressing Status and Correlation of ARID1A and Histone H2B on Breast Cancer. *Biomed Res. Int.* **2016**, 1–9.
117. von Minckwitz, G. *et al.* (2018) Trastuzumab Emtansine for Residual Invasive HER2-Positive Breast Cancer. *N. Engl. J. Med.* **380**, 617–628.

APPENDIX

APPENDIX I

Breast cancer TNM staging

Supplementary Table 1. TNM classification of BrC according to AJCC/UICC guidelines (adapted from²⁷).

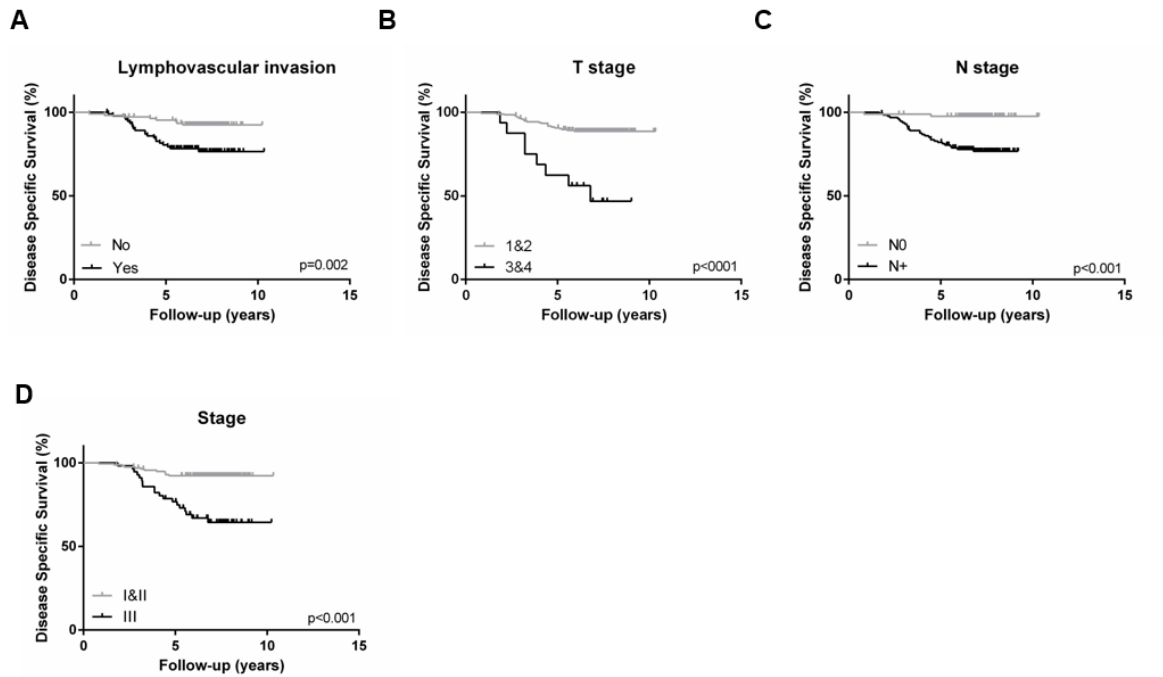
T- Primary tumour (Clinical and pathological)	
TX	Primary tumour cannot be assessed
T0	No evidence of primary tumour
Tis (DCIS)	Ductal carcinoma <i>in situ</i>
T1	Tumour ≤ 20 mm in greatest dimension
T2	Tumour >20 and ≤ 50 mm in greatest dimension
T3	Tumour > 50 in greatest dimension
T4	Tumour of any size with direct extension to the chest wall and/or to the skin; Inflammatory carcinoma
N – Regional Lymph Nodes (Clinical and Pathological)	
NX	Regional lymph nodes cannot be assessed
N0	No regional node metastases
N1	cN1* Metastasis to movable ipsilateral Level I, II axillary lymph node(s)
	pN1** Micrometastasis or macrometastasis in 1-3 axillary lymph nodes
N2	cN2 Metastasis in ipsilateral level I, II axillary lymph nodes that are clinically fixed; or matted or ipsilateral internal mammary nodes in the absence of axillary lymph node metastasis
	pN2 Metastasis in 4-9 axillary lymph nodes; or positive ipsilateral internal mammary lymph nodes by imaging in the absence of axillary node metastasis
N3	cN3 Metastasis in ipsilateral infraclavicular lymph node(s), ipsilateral internal mammary lymph node metastasis or metastasis in ipsilateral supraclavicular lymph node(s)
	pN3 Metastasis in 10 or more axillary lymph nodes or in infraclavicular lymph nodes, positive ipsilateral internal mammary lymph nodes by imaging in the presence of one or more positive Level I, II axillary nodes, more than 3 axillary lymph node metastases by sentinel lymph node biopsy in clinically negative ipsilateral internal mammary lymph nodes or in ipsilateral supraclavicular lymph nodes
M – Distant Metastasis (Clinical and Pathological)	
M0	No clinical or radiographic evidence of distant metastasis
M1	Distant metastasis detected by clinical and radiographic means and/or histologically proven metastasis larger than 0.2mm

APPENDIX II

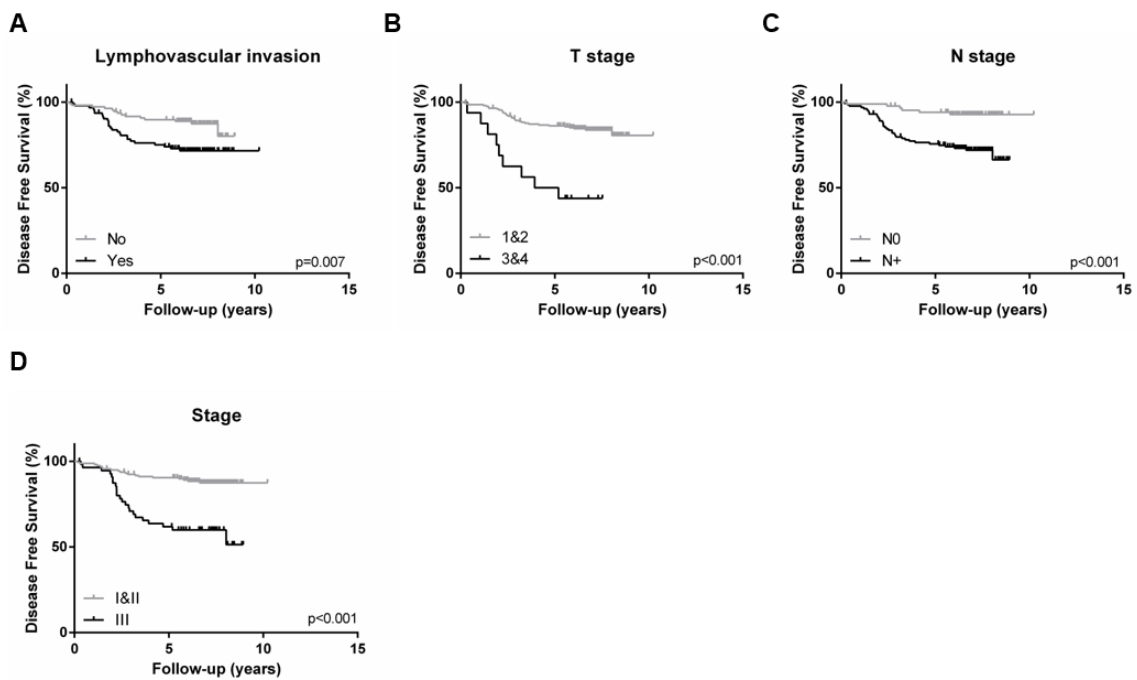
Supplementary Table 2. Anatomic Stage/Prognostic Groups (adapted from²⁷).

Anatomic stage	Prognostic groups		
	T	N	M
0	is	0	0
IA	1	0	0
IB	0	1mi	0
	1	1mi	0
IIA	0	1	0
	1	1	0
	2	0	0
IIB	2	1	0
	3	0	0
IIIA	0, 1, 2	2	0
	3	1, 2	0
	4	0, 1, 2	0
IIIB	4	0, 1, 2	0
IIIC	Any	3	0
IV	Any	Any	1

APPENDIX III



Supplementary Figure 1. Disease-specific survival curves (Kaplan-Meier with log rank test) of clinicopathological parameters.



Supplementary Figure 2. Disease-free survival curves (Kaplan-Meier with log rank test) of clinicopathological parameters.