The Lancet Oncology

Defining genomic, transcriptomic, proteomic, epigenetic, and phenotypic biomarkers with prognostic capability in male breast cancer: a systematic review

--Manuscript Draft--

Manuscript Number:	THELANCETONCOLOGY-D-22-01014R2
Article Type:	Review (Post author-enquiry)
Keywords:	male breast cancer; biomarkers; translational medicine; Systematic review
Corresponding Author:	Valerie Speirs University of Aberdeen Aberdeen, UNITED KINGDOM
First Author:	Subarnarekha Chatterji, MSc
Order of Authors:	Subarnarekha Chatterji, MSc
	Emma Krzoska, MChem
	Christopher Thoroughgood, PhD
	John Saganty, MPhil
	Peng Liu, PhD
	Beatrix Elsberger, PhD
	Rasha Abu-Eid, PhD
	Valerie Speirs, PhD
Manuscript Region of Origin:	UNITED KINGDOM
Abstract:	While similar phenotypically, there is evidence that male and female breast cancer differ in their molecular landscapes. In this systematic review, we consolidated all existing prognostic biomarker data in male breast cancer, spanning genetics, transcriptomics, proteomics, and epigenetics as well as phenotypic features of prognostic value from articles published in a 29-year period (1992 – 2021). We identified knowledge gaps in the existing literature, discussed limitations of included studies, and outlined potential approaches for translational biomarker discovery and validation in male breast cancer. We also recognised STC2, DDX3, and DACH1 as underexploited markers of male-specific prognostic value in breast cancer. Finally, beyond describing the cumulative knowledge on the extensively researched markers ERα, PR, HER2, AR, and BRCA2, we highlighted ATM, CCND1, FGFR2, GATA3, HIF1α, MDM2, p53 and c-Myc as well-studied predictors of poor survival, that also aligned with several hallmarks of cancer.

Title: Defining genomic, transcriptomic, proteomic, epigenetic, and phenotypic biomarkers with prognostic capability in male breast cancer: a systematic review

Authors:

[1,2]Subarnarekha Chatterji, MSc (ORCID: 0000-0002-8980-4982)

[1†]Emma Krzoska, MChem (ORCID: 0000-0002-2755-6485)

[1†]Christopher W Thoroughgood, PhD (ORCID: 0000-0002-4470-5411)

[1†]John Saganty, MPhil (ORCID: 0000-0002-3523-227X)

[1,2]Peng Liu, PhD (ORCID: 0000-0002-0058-0434)

[3]Beatrix Elsberger, PhD (ORCID: 0000-0002-2864-5789)

[1,4]Rasha Abu-Eid, PhD (ORCID: 0000-0002-6634-0329)

[1,2*]Valerie Speirs, PhD (ORCID: 0000-0002-0602-4666)

[†]*Authors contributed equally*

Affiliations:

1. School of Medicine, Medical Sciences, and Nutrition, University of Aberdeen, Aberdeen AB25 2ZD, UK
2. Institute of Medical Sciences, University of Aberdeen, Aberdeen AB25 2ZD, UK
3. NHS Grampian, Aberdeen Royal Infirmary, Breast Unit, Foresterhill Rd, Aberdeen AB25 2ZN, UK
4. Institute of Dentistry, University of Aberdeen, Aberdeen AB25 2ZR, UK

*****Full Professor and corresponding author: Professor Valerie Speirs

(valerie.speirs@abdn.ac.uk)

Institutional corresponding address:

Institute of Medical Sciences

University of Aberdeen

Foresterhill

Aberdeen AB25 2ZD

United Kingdom

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

All authors agree with the content of the manuscript and consent to publication.

Availability of data and materials

Not applicable. All information can be found in the Appendix and the reference list.

Competing interests

VS received funding from the University of Aberdeen Development Trust. The other authors declared no conflicts of interest.

Funding

This work was supported by a University of Aberdeen Development Trust Elphinstone Scholarship (SC) and NHS Grampian Endowments (VS).

Author's contributions

Study concept and design: SC, RAE, and VS; Literature search, title screening, and abstract screening: SC; Full-text screening and data extraction: SC, EK, CT, JS, and PL; Accuracy checks: PL, RAE, and VS; Writing – original draft: SC; Writing – review and editing: SC, RAE, BE and VS; Supervision: RAE and VS.

EK, CT, and JS contributed equally. All authors approved the final version to be published.

None of the authors are employed by NIH.

Abstract

While similar phenotypically, there is evidence that male and female breast cancer differ in their molecular landscapes. In this systematic review, we consolidated all existing prognostic biomarker data in male breast cancer, spanning genetics, transcriptomics, proteomics, and epigenetics as well as phenotypic features of prognostic value from articles published in a 29-year period (1992 – 2021). We identified knowledge gaps in the existing literature, discussed limitations of included studies, and outlined potential approaches for translational biomarker discovery and validation in male breast cancer. We also recognised STC2, DDX3, and DACH1 as underexploited markers of male-specific prognostic value in breast cancer. Finally, beyond describing the cumulative knowledge on the extensively researched markers ERα, PR, HER2, AR, and *BRCA2*, we highlighted *ATM*, *CCND1*, *FGFR2*, *GATA3*, HIF1α, *MDM2*, p53 and c-Myc as well-studied predictors of poor survival, that also aligned with several hallmarks of cancer.

Introduction

Breast cancer (BC) affects both sexes but is around 100 times rarer in men[1]. Latest statistics from 2019 show that 25,143 men were affected worldwide, with a 48.1% mortality rate[2]. In comparison, BC affected 1,977,212 women during this period with 34.8% mortality rate[2]. Current clinical management of male breast cancer (MBC) is identical to female breast cancer (FBC), informed by female-only clinical trials. However, MBC differs from FBC in clinical presentation, distribution of histopathological types, and hormone receptor (HR) expression[1,3-5]. Clinical presentation is typically late, MBCs are predominantly oestrogen receptor (ERα) positive (up to 95%), with human epidermal growth factor receptor 2 (HER2) expression uncommon, and triple negativity extremely rare in men[4,6-9].

Hierarchical clustering studies on genetic, transcriptomic, and epigenetic data have identified MBC-specific clusters of prognostic value with limited overlap with the Prediction Analysis of Microarray 50 (PAM50) intrinsic subtypes in FBC[10-15]. Germline mutations in *BRCA2*, established as a high penetrance MBC susceptibility gene have also been extensively researched. Carriers have a lifetime risk of up to 10% of developing cancer, frequently with poor prognosis and aggressive disease characteristics[16-19]. However, despite growing consensus on high-risk men with relevant family history to be offered screening, such an initiative does not yet exist.

Biomarker studies in MBC are few despite rising interest over the past decade. Large scale collaborative studies like the International Male Breast Cancer Program have concentrated mainly on ERα, PR and HER2, which are already integrated into clinical practice[7]. Novel biomarker studies in MBC have revealed numerous candidates with possible male-specific value, but most suffer from small cohorts and lack of independent validation, meaning these remain under-investigated.

While many general reviews on MBC exist, to our knowledge there is no comprehensive systematic review to identify knowledge gaps in MBC biomarkers with prognostic potential.

Hence, we exhaustively reviewed molecular studies in MBC adopting a multi-omics and phenotypic approach. We comprehensively describe the existing landscape of prognostic biomarkers in MBC and highlight several molecules that could provide complementary information beyond what is established in BC for future clinical management.

Methods

We conducted and reported this systematic review following Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) recommendations[20].

Search strategy and selection criteria

A systematic search of published literature on MBC biomarkers with a multi-omics and phenotypic approach was conducted using PubMed, Medline, Scopus, Embase, and Web of Science, from the inception of the databases to 16th June 2020. An updated search was performed between 17th June 2020 and 1st November 2021 to include the most recent publications. The representative terms “TITLE (male OR men) AND TITLE (breast OR mammary OR “mammary gland”) AND TITLE (neoplasm OR neoplasia OR malignancy OR malignancies OR cancer OR carcinoma OR tumour OR tumor) AND (KEY (biomarker OR marker)) were used to conduct the electronic search. Complete database specific search terms are detailed in the Appendix (Page 3).

Inclusion criteria were:

- Primary study population must have included MBC patients and should have been the focus of the study
- Studies must have investigated marker(s) of any omics type or morphological and/or phenotypic features with respect to disease pathogenesis/progression/survival and clinicopathological characteristics of study population(s)

Exclusion criteria were:

- Case reports, case series, letters to the editor, conference abstracts, comments, reviews, and systematic reviews
- Studies conducted on species other than humans
- Original articles in languages other than English
- Primary cohort size ≤ 5

No restrictions were made on methodology, statistical significance of results, or diagnostic/prognostic/predictive value of the biomarkers studied. The selection criteria were intentionally broad to ensure exhaustivity and minimize loss of information. Additionally, reference lists of the included manuscripts were manually searched by SC to identify studies that may have been missed by the electronic search.

Abstracts retrieved from these searches were exported to EndNote referencing software, using which deduplication and screening of titles and abstracts to exclude studies that did not fulfil inclusion criteria was done by SC. Full-text screening of the short-listed articles was conducted in pairs by SC, EK, CT, JS, and PL.

Data extraction

Data extraction of the following variables was performed using Microsoft Excel: first author, published year, country/countries where the study was conducted, study design, method(s), type of tissue tested, cohort size, control group, age (mean/median and range), anatomic stage, histological type and grade, treatment information, St. Gallen classification, nodal status, HR (ERα, PR, HER2) status, number of biomarkers studied, biomarker type (prognostic/predictive/diagnostic), biomarker category (genetic/transcriptomic/proteomic/epigenetic/phenotypic), survival associations, and associations with clinical features described in each article.. FBC data were recorded using the same criteria when present and relevant. To ensure uniformity, all reviewers extracted data from five randomly selected articles for training and calibration. For articles identified in the original search conducted on 16th June 2020, the data extraction process was conducted

by two independent reviewers in three pairs (SC + EK, SC + CT, SC + JS). Disagreements were resolved through discussion and with the involvement of a third reviewer when necessary. Data extraction for articles identified in the search from 17th June 2020 to 1st November 2021 was done following the same protocol by SC and PL. Accuracy checks were performed on at least 10% randomly selected articles by RAE and VS.

Quality assessment

Risk of bias assessment was conducted using the Joanna Briggs Institute Critical Appraisal tools using checklists for case-control studies, and analytical cross-sectional studies, as appropriate[21]. Studies had high risk of bias if the response to at least one appraisal question was “No” and/or to multiple questions was “Unclear”. If one question had an “Unclear” response, but all other responses were “Yes”, the risk of bias was moderate. If the response to all questions was “Yes”, the risk of bias was low.

Results

Database search results

In total, 1359 records were retrieved from 5 databases: 306 (PubMed), 576 (Scopus), 187 (Medline), 158 (Embase), 132 (Web of Science). Duplicates (682) were removed, following which 677 articles were screened based on title and abstract. Then, 480 articles were removed as they did not meet the inclusion criteria, leaving 197 articles. These underwent full-text screening, after which 20 articles were removed for not fulfilling the inclusion criteria. Data extraction was performed on the remaining 177 articles. A manual reference search within included articles revealed 20 relevant articles that were missed by the electronic search. In total, 197 articles were finally included. A PRISMA chart is shown in the Appendix (Page 126).

The included studies were conducted from1992 to 2021. Of these, 27 were descriptive[22-48], and 35 were screening studies[49-82]. Of the latter, 26 reported mutations without any clinical associations. 64 studies[6,7,11,13,15,83-141] reported biomarkers linked to survival and the remaining 78 studies reported biomarkers with clinical associations[10,12-14,49,54,58,59,61,66-68,71,84,87,90,93,106,142-209].

Study Characteristics

We identified 76 case-control studies[10,13-15,22-24,26,31,34,38,44,48,51,54,55,58,60,67,72,73,81,84,88,90,92,94,98,99,106,109,110,120,123,132,136,142-146,148-150,152-160,162,164,166,167,171,172,175-177,182-185,188,191,193,196,202,203,205,206,209,210], of which MBC outcomes were measured against gynaecomastia in 10 studies[23,34,106,132,153,154,158,172,177,193], FBC data in 43 studies[10,13-15,26,31,34,38,48,73,81,84,88,90,92,98,99,109,110,120,123,136,155,159,160,162,164,167,171,175-177,182-185,188,191,196,202,206,209,210], healthy men, women or both in 23 studies[22,24,51,54,55,58,60,67,72,94,142-146,148-150,152,157,175,203,205], and 1st degree male relatives with history of cancer (non-breast) in 1 study[166]. Normal male breast tissue[10,15,44,132,156,162,209], lymph node tissue[156], and non-malignant breast cell lines[10] were used as controls in 7, 1, and 1 study, respectively. Of the case-control studies, 80.3% (n = 61), 5.3% (n = 4), and 14.4% (n = 11) articles had high, moderate, and low risk of bias, respectively (Appendix Page 5).

The remaining 121 studies were cross-sectional[6,7,11,12,25,27-30,32,33,35-37,39-43,45-47,49,50,52,53,56,57,59,61-66,68-71,74-80,82,83,85-87,89,91,93,95-97,100-105,107,108,111-119,121,122,124-128,130,131,133-135,137-141,147,161,163,165,168-170,173,174,178-181,186,187,189,190,192,194,195,197-201,204,207,208,211,212]. Most had MBC patients as their sole cohort, while 2 studies included FBC patients with MBC-affected relatives alongside their primary MBC cohort[76,79]. Out of the cross-sectional studies, 56.2% (n = 68) and 43.8% (n = 53) had high and low risks of bias, respectively (Appendix Page 11). Study characteristics are summarized in the Appendix (Page 19).

We identified 304 biomarkers in total and classified them according to their respective omics/phenotypic categories. The 10 most studied biomarkers from each category, based on

the number of reporting studies and associations with clinical features are detailed in Tables 1-4. The full list of biomarkers with their clinical associations, and all reported pathological gene variations are provided in the Appendix (Page 43-125).

Proteomic markers

ERα, PR, and HER2

These receptors currently define standard-of-care in BC and were studied both as biomarkers and clinical factors associated with other biomarkers. The MBC cohorts studied were overwhelmingly ERα-positive, predicting improved OS and DFS[7,123], while ERα-negativity, predicted reduced OS[104,118,122,134] and younger age of diagnosis[93]. Like FBCs, PR was frequently co-expressed with ERα, its positivity mostly predicting prognostic benefit[7,87,93,104,105,118,122].

Overexpression and amplification of HER2 was evaluated by immunohistochemistry (IHC) and fluorescent *in-situ* hybridisation (FISH), the latter being detailed in the genetics/transcriptomics markers section. Overexpression was associated with aggressive features and reduced survival by every study investigating HER2 prognostic value[6,87,95,101,129,188,198] (Table 1).

St Gallen surrogate classification

Luminal B and triple negative MBCs had poor survival and aggressive features[87,101,119,190,208], with the latter more frequent in men of black ethnicity[101]. Basal-like MBCs were diagnosed at younger age than Luminal A/B MBCs[190]. Several biomarkers were expressed differentially between the Luminal classifications. GCDFP15-positivity[187] and p53-negativity[181] were associated with Luminal A MBCs, while ATF3, FATP1, p21-positivity, and Bcl2-negativity were associated with HER2-negative Luminal B MBCs[93,100]. The latter also had higher expression of EGFR and NF-κB compared to Luminal A MBCs[37] (Appendix Page 41).

Other proteomic markers

AR expression had both prognostic advantage[6,7,116,123,131,179,200] and disadvantage[94,96,117]. Interestingly, two out of three studies predicting poor outcome were conducted on ethnically homogeneous Chinese populations[94,117]. Like FBCs, AR was consistently co-expressed with ERα[94,116,131,133,179]. AR co-expression with ERα and FOXA1 predicted improved OS[123] and DFS[6], respectively.

High tumour proliferation index (represented by Ki-67/MIB1 index) consistently predicted poor survival and aggressive disease[87,93,113,115,118,129,131,133,135,184,186,196,197].

Of the most studied markers, p53[93,119,128,129,131], p21[93,125,160,196], EGFR[118,188,190] and c-Myc[125,129] predicted reduced survival. The tumour hypoxia markers HIF1-α, CA-9 and Glut-1 along with their co-expression profiles also predicted poor outcome[124,141,180].

Relatively few biomarkers predicted improved outcome and were rarely reported by multiple studies. Bcl-2[93,181,189,194,202] and Cyclin D1 positivity[93,121,125,133] were mostly linked to improved outcome.

Several markers displayed sex-specific differences in expression. Hormone receptors ERα[185], PR[202], AR[123], ERβ1[123] and ERβ2[123] were expressed more frequently in MBCs than FBCs. STC2[109], IGF1-R[188], CAXII[188], p21[160,196], p27[196], p53[160] and Bcl-2[202] were also overexpressed in MBC compared to FBC, while the opposite was true for DACH1[182], PD-1[183], MET[188], FGFR2[188], CD44v6[188] and GATA3[120]. DDX3 had higher cytoplasmic expression but lower nuclear expression in MBCs compared to FBCs[102]. Improved survival or favourable outcomes in MBC were linked to STC2[109], p27[125,196,197], Bcl-2[93,181,189], and high cytoplasmic DDX3 expression[102]. The opposite was true for p21[93,125], p53[31,93,119,128,129,131,160,181,202], DACH1[182], and GATA3[90,120]. The prognostic value of STC2[109], DDX3[102], and DACH1[182] were assessed by only one study each (Table 1 and Appendix Page 43).

Genetic and transcriptomic markers

Pathogenic variations in BRCA genes with prognostic value

Germline *BRCA2* mutations are the most frequently reported pathological gene variations in MBC. These predicted reduced overall (OS), disease-free (DFS), and disease-specific survival (DSS)[85,87,96], and aggressive features like young age of diagnosis, bilaterality, contralaterality, node positivity, advanced tumour grade, ERα/PR-negativity, HER2-positivity, high Ki-67 index, personal history of cancer[59,61,68,87,149,164,167,170,173,175], high frequency of genetic aberrations[175], amplifications[88] and copy number variations (CNV)[168] of several cancer-related genes. *BRCA2* mutations were more frequent and had more aggressive features in MBCs compared to FBCs[59,164]. In contrast, germline *BRCA1* mutations were less frequent in MBCs[59] and had less pronounced prognostic value, with links to advanced tumour grade[164], ERα-negativity[170], and family history of pancreatic cancer[66] (Table 2).

Germline mutations were most frequently reported in *BRCA2* and *BRCA1* (28 and 12 studies, respectively), followed by *CHEK2, PALB2*, and *ATM* (9, 7, and 3 studies respectively).

Pathogenic variations in other genes with prognostic value

While uncommon in MBC (0 - 9% of all cases[6,7,123]), HER2 amplification predicted reduced OS, younger age of diagnosis, large tumour size, advanced disease stage, and both regional and distant metastasis[84,86,93,95].

Several genetic variations predicted reduced OS. These included somatic mutations in *PIK3CA*[88], *GATA3*[90] and *THY1*[92], and amplifications in *MDM2, PAK1, TGFB2, SCYL3*[88], *CCND1* and *EMSY*[84]. Mutations in DNA repair genes were enriched in Luminal A-like MBCs compared to matched FBCs and predicted reduced survival in general[90]. In contrast, survival benefit was associated with relatively few genetic/transcriptomic variations, with only upregulation of miR-125b, which targets genes covering multiple biological signalling pathways in many cancers[213], being reported in >1 study[177,209] (Table 2 and Appendix Page 71).

Pathogenic variations associated with MBC risk

Germline mutations in *PALB2* and *RAD51D*[54] had the highest odds-ratios (17.30, 8.58; 11.20, 10.18 , using the Exome Variant Server and Non-Finnish European datasets, respectively), followed by *MUTYH* (4.54)[147], *CHEK2* (4.47)[58], and *SULT1A1* (3.09; A/A polymorphism)[148]. Copy number (CN) gain in *PALB2* was associated with node negativity[12] and its mutated status was associated with bilaterality[49]. Increased MBC risk was also linked to single nucleotide polymorphisms (SNPs) in multiple genes, with rs3803662 (*TOX3*) reported by two independent groups[144,145].

Screening studies from 1995 to 2021 identified pathogenic mutations in several genes in MBC, most of them germline. The *CHEK2* c.1100delC mutation was reported most frequently[49,52,54,58,63,66], followed by the *BRCA2* c.6174delT[57,61,64,66] and c.771_775delTCAAA (also known as c.999del5)[59,69,72,81] (Appendix Page 100).

Epigenetic markers

Advanced tumour grade, high mitotic index, large tumour size, ERα-negativity, and mutated *BRCA2* were linked to promoter hypermethylation of most reported genes[83,155,156]. Interestingly, hypermethylated *RASSF1A* and *RARB* were linked to both ERα-negativity and PR-positivity, which have opposing clinical significance in FBC[157]. Hypermethylated *RASSF1A* was also linked to HER2-positivity[156]. High methylation indices, high methylation rate, and high number of methylated genes predicted reduced OS and DSS, and aggressive features like *BRCA2*-mutation, high mitotic index, high tumour grade, and large tumour size[15,83]. Only one study associated promoter hypermethylation of any gene to survival, with hypermethylated *TWIST1* predicting reduced DSS, especially in *BRCA2*-mutated MBCs[83].

Conflicting results were reported on *AR* promoter hypermethylation. Virtually non-existent *AR* methylation and very little methylation of its co-regulators was observed in MBC when compared to gynaecomastia [154]. However, tumour DNA had higher *AR* methylation compared to normal tissue and lymph nodes (both patient unmatched)[156]. *AR* hypermethylation was also associated with wild type *BRCA1/2*[156].

Regarding sex-specific epigenetic differences, reduced methylation levels were more common in both invasive carcinoma (IC) and ductal carcinoma *in-situ* adjacent to invasive carcinoma (DCIS-AIC) in MBC compared to FBC. Only *GATA5, THBS1, MSH6*, and *RASSF1A* were more heavily methylated in males compared to females[155,157].

Within MBC cohorts, higher methylation was reported in DCIS-AIC compared to pure ductal carcinoma *in-situ* (DCIS), while IC had higher methylation levels compared to DCIS-AIC. Hypermethylation in normal breast tissue and lymph nodes (both patient unmatched) was consistently less frequent compared to IC[156] (Table 3 and Appendix Page 113).

Morphological and/or phenotypic features

Several morphological features of MBC had prognostic significance. Unsurprisingly, high mitotic activity index predicted reduced survival[137]. High nuclear area and high variation in nuclear size predicted poor survival and aggressive features [128,138]. Presence of fibrotic foci predicted reduced OS[124,137] and recurrence-free survival (RFS)[137], and advanced tumour grade, nodal involvement, and low tubule formation[124]. The latter also predicted reduced OS[138]. Like FBCs, low density of tumour infiltrating lymphocytes (TILs) predicted reduced OS and RFS[137], and nodal involvement[186]. Intriguingly, HER2-positive MBCs had higher density of TILs than HER2-negative MBCs, although HER2 overexpression predicted poor prognosis[137].

Low grade ERα-positive MBCs had reduced elastosis than matched FBCs. In FBCs elastosis is strongly associated with ERα expression. Therefore, low frequency of elastosis in MBC despite overwhelming ERα-positivity suggests sex-specific ERα action[206].
Morphological features of both lymphangiogenesis and angiogenesis like high lymphatic vessel density, high distribution of lymphatic vessels, and high frequency of vascular invasion were linked to advanced tumour grade, high tumour proliferation index, and hormone receptor negativity, albeit without reproduction[186]. In agreement, high CD34

expression representing microvascular density predicted reduced RFS and advanced disease stage[130] (Table 4 and Appendix Page 119).

Novel subgroups in MBC

The first major hierarchical clustering study identifying male-specific BC subgroups was done by Johansson et al[13]. Luminal M1 group exhibited HER2-positivity and associated with invasion, proliferation, and metastasis, while Luminal M2 group displayed ERα-positivity and associated with anti-tumour immune response[13]. They also previously identified Male-simple and Male-complex clusters. The former was genetically stable and differed from female intrinsic subtypes, while the latter consisted of *BRCA2*-mutated MBCs, with worse prognosis and genetic overlap with the Luminal B intrinsic type[14].

These results were validated by a genome-wide methylation study revealing two stable MBC epitypes (ME1 and ME2)[10]. ME1 epitype displayed high mitotic activity, high fraction of genome alteration, Cyclin A-positivity, and ERα-negativity, and frequent hypermethylation of genes involved in key pathways (H3K27me3 epigenetic silencing, transcriptional regulation with HOX genes, WNT, TGF-β, and MAPK signalling, cellular and focal adhesion, and FGFR ligand binding and activation). ME1 and ME2 epitypes aligned with the Luminal M1 and M2 subgroups, respectively[13].

A later study reported 4 epigenetics-based clusters based on the relative promoter hypermethylation levels of *RASSF1A, GSTP1, WIF1, RARB,* and *MAL*. Notably, Cluster 3 associated with mutated *BRCA2* ($p = 0.02$)[83]. This study performed a subgroup analysis on *BRCA2*-mutated MBCs which separated into 2 clusters based on the hypermethylation levels of *GSTP1, MAL,* and *RASSF1A*[83].

Most recently, two clusters were reported based on RNASeq data[11]. Cluster 1 had reduced OS and associated with HER2 signalling, proliferation, invasion and metastasis, and immune response, while Cluster 2 associated with the apoptosis hallmark and NAT1 signalling[11]. These clusters had limited overlap with the Luminal M1 and M2 subgroups. Immune

response clustered with invasion and metastasis, and proliferation, directly contradicting Luminal M1 and M2 characteristics[11,13].

Cluster separation was also reported based on chromosome 16q CNVs. Cluster A had low rates of CN gain and amplification, predicting prognostic benefit, while Cluster B had aggressive features[84]. Building on this work, another study reported clusters based on chromosome 16q CNVs, where Cluster A associated with node positivity, and Cluster B with triple negativity[12].

Four clusters based on immunohistochemical markers were described[93]. Clusters A1 and A2 had aggressive characteristics; A1 defined by hormone negativity, and A2 by ERα-positivity, PR-negativity, and HER2-amplification. The less aggressive clusters B1 and B2 were histologically identical, although B1 exhibited BRST-2 positivity and nodal involvement, while B2 had the opposite features[93].

MBC clusters separating on ER/PR isoforms were also reported[123]. These respectively separated on the cytoplasmic expression of ERβ1 and 2, PR isoforms A and B, and collective action of AR with ERα and β1 isoforms. Only cytoplasmic-ERβ cluster had FBC overlap[123] (Table 5).

Alignment of biomarkers with the Hallmarks of Cancer

Upon interrogation of the COSMIC database[214], certain genetic, transcriptomic, proteomic, or epigenetic markers aligned with the 2000 and 2011 Hallmarks of Cancer[215,216]. These had prognostic impact in MBC and/or differential expression between the sexes. Certain molecules identified in the same categories were also speculatively linked to the most recent Hallmarks of Cancer[217] (both described on page 127 of the Appendix). Based on these associations, these molecules may warrant further research: *ATM*, *CCND1* (Cyclin D1), *GATA3*, *FGFR2*, *HIF1A* (HIF1-α), *MDM2*, *MYC* (c-Myc), and *TP53* (p53). These were linked to multiple hallmarks of cancer through promoter and/or suppressor action, were associated

with ≥1 clinical feature across multiple omics categories and could predict survival in at least one of these categories.

Discussion

MBC is receiving increased recognition. A bibliometric analysis revealed that most publications in MBC focused on clinical risk factors and management, followed by comparisons against FBC[218]. MBC management is still largely defined by superficial extrapolation of FBC standard-of-care despite mounting evidence of sex-related differences. Recognising a need to identify translationally valuable biomarkers that can define a male-inclusive picture of BC, this systematic review comprehensively described the biomarker landscape of MBC and identified markers that may aid future clinical management. To our knowledge, this is the first exhaustive systematic review on the subject.

ERα and PR emerged as having sex-specific regulatory characteristics. Although a known modulator of ERα binding in FBC, many PR binding sites were devoid of ERα in MBC[98]. Hierarchical clustering studies found independent PR clusters[123] in MBC, while ERα/PR action clustered together in FBC[98,123]. Mathematical modelling revealed no continuous dependency effect on ERα for PR[31]. Furthermore, two FBC clusters were identified based on PR action in FBC but not in MBC[171].

Regarding ER isoforms, ERα/ERβ/AR[123], and ERα/FOXA1/AR coaction predicted improved survival in MBC[6]. As most ERα binding sites in both sexes are independent of FOXA1[98], this suggests an intermediary role of FOXA1 (and possibly ERβ) in ERα/AR interaction in MBC. This requires elucidation.

AR expression, when studied independently, predicted contradicting prognostic outcomes[6,7,94,96,116,117,123,131,179,200]. Epigenetic findings on AR were also inconsistent. AR hyperactivity in ERα-positive MBC was speculated based on hypomethylation of AR and its co-regulators compared to gynaecomastia[154], while another study demonstrated AR

hypermethylation in tumours compared to unmatched normal lymph nodes and breast tissue[156]. Therefore, the exact impact of AR methylation remains unclear. The contradictory role of AR was further highlighted by its value as a therapeutic target in MBC. Phase II trial data showed that the AR inhibitor enzalutamide was well-tolerated in both sexes, and improved PFS in both HR positive and androgen-driven triple negative BC[219,220]. Similar results were seen with the AR/CYP17-L inhibitor seviteronel in both sexes[221]. In FBC, AR plays a compensatory role for ERα in ERα-negative/AR-positive FBC, and this is supported by overlapping binding characteristics of ERα and AR[98,222]. However, the same cannot be speculated for MBC as most patients are ERα/AR-positive. A partial explanation is offered by the sex-specific nature of prognostic ability of ERα binding sites[98], but we await a complete picture of ERα/AR interaction in MBC. Intriguingly, AR-driven tumour-suppressor activity was observed in ERα/AR-positive BC cell lines and FBC patient-derived explant (PDE) models, clearly supporting agonism over antagonism of AR as a more valuable treatment strategy[223].

The aggressive nature of germline *BRCA2* mutations has been established in MBC [59,61,68,87,149,164,167,170,173,175]. However, *BRCA2* is yet to inform clinical management, despite there being an argument for male patients with family history of *BRCA2*-related cancers (breast, ovarian, prostate, and pancreatic) to be screened and offered genetic counselling[224]. The incidence of *BRCA2*-mutated MBCs in different ethnicities also need to be established.

Given the negative prognostic effect of somatic mutations in the *PIK3CA* gene in MBC[88,158], the SOLAR-1 trial is worth mentioning. This randomised phase-3 trial included men and postmenopausal women with HR-positive/HER2-negative BC with mutated *PIK3CA* and demonstrated improved OS when the PI3KA-specific inhibitor Alpelisib was administered with Fulvestrant[225]. This trial is an encouraging example of positive advances being made towards inclusion of men in clinical trials.

Discovery of novel markers in MBC has historically suffered due to small cohort sizes and lack of prospective validation. This generally aligns with the broader picture of biomarker

discovery in oncology, where most molecules are often left unexplored beyond their initial identification and establishment of a significant survival association. The relative rarity of MBC and small number of research papers brings this into sharp focus.

As shown in the Appendix (Page 129), most of the well-studied biomarkers with hallmarks functions also regulate the G1/S phase transition pathway of the cell cycle along with *RB1, MDM2, ATR, CHEK2, CDKN1A* (p21)*, CDKN1B* (p27)*, CDKN2A,* and *CCNE1*, alterations of which were also linked with MBC clinical outcome in at least one -omics category (Figure 1). Most of these biomarkers predicted poor survival, which justifies focused drug-target identification studies through selective inhibition of regulatory pathways. The role of Cyclin D1 is especially worth investigating, as it predicted improved survival as a proteomic marker[93,121,125,133], but the opposite as a genetic marker (*CCND1*)[84].

In this regard, the CDK4/6 inhibitor Palbociclib was approved for use in metastatic MBC[226]. Literature supporting the use of CDK4/6 inhibitors in combination with tamoxifen/AI and GnRH in a metastatic setting also exist[227,228]. A recent case report described complete remission of a metastatic MBC patient following treatment with Abemaciclib, Fulvestrant, and Leuprolide[229]. The evidence gathered here supports this approach. However, extending this to the adjuvant setting for MBC may be premature based on results of the PALLAS trial[230].

Amongst the plethora of molecules we identified, STC2[109], DDX3[102], and DACH1[182] are especially worth highlighting in those that were only reported in single studies. STC2 is involved in pathways regulating stress response, hypoxia, apoptosis prevention, cellular proliferation, migration, and immune response[231]. Tumour and stromal STC2 expression were observed in some 50% and 65% of MBC patients, respectively[109]. DDX3 promotes cancer progression by remodelling the tumour microenvironment[232]. Nuclear and cytoplasmic expression of DDX3 was observed in 42.5% and 20.8% of MBC patients, respectively[102]. DACH1 is a tumour suppressor implicated in the inhibition of invasion and metastasis via downregulation of matrix metalloproteinase 9 transcription, whose positivity was observed in 35.7% MBC cases[182,233]. These proteins were differentially expressed

between the sexes and could predict survival in MBC, however, remains underexploited from a translational perspective.

Defining morphological markers of prognosis is necessary as these can be the primary diagnostic considerations. Variation in nuclear area and size are obvious markers of negative prognosis in MBC, which was confirmed in two studies we reviewed[128,138]. The presence/dimensions of fibrotic foci emerged as important markers predicting reduced survival[124,137]. Suggested to be the link between hypoxia and aggressive tumour characteristics, these results were validated by the unfavourable prognostic value of the hypoxia markers HIF1-α, CA-9, and Glut-1[124,141].

Ethnic homogeneity may explain lack of reproducibility for certain studies, such as conflicting prognostic impact for certain markers. This is concerning, as US data show that the age-standardized incidence of MBC in non-Hispanic black men is 2.6 times higher than their white counterparts for ERα-positive/HER2-negative BC[234]. Despite this, no molecular studies investigating ethnicity-specific differences in MBC exist, leaving a significant knowledge gap. Also, ethnicities were not specified in the clustering studies, and therefore no conclusions could be drawn regarding their global representation.

The appropriate selection of controls is another area that may require future consideration. For example, some studies used gynaecomastia samples as controls, as normal male breast tissue is difficult to obtain. However, gynaecomastia is now treated as being aetiologically distinct from MBC and therefore unlikely to be a suitable comparison[235,236] presenting potential limitations.

Conclusion

Our results demonstrate MBC is a heterogeneous and complex condition with striking distinctions from FBC. MBC research has seen remarkable evolution, from simply replicating

FBC marker studies, to its treatment as a separate condition with exploratory studies contributing to a male-specific molecular profile.

We identified conflicting evidence regarding regulation, expression, and prognostic utility of key BC markers alongside sex-specific differences. Considering this, the role of ERα, PR, and AR need to be re-established in a male-specific setting. Developing suitable MBC laboratory models are necessary to achieve this. Beyond the established BC markers, we highlighted that STC2, DDX3, and DACH1 may have grounds for further investigation. We also identified *ATM*, *CCND1* (Cyclin D1), *FGFR2*, *GATA3*, *HIF1A* (HIF1-α), *MDM2*, *MYC* (c-Myc) as well studied predictors of poor prognosis.

To effectively drive the inclusion of male-specific biomarkers from bench to clinical practice, inclusion of men in randomized clinical trials is crucial. Positive advances have been made in this respect with the International Male Breast Cancer Program making a concerted effort to run male-specific trials, and at least two MBC phase-II trials investigating GnRH/AI/tamoxifen and AR-antagonists being reported[221,237,238] alongside the SOLAR-1 trial discussed above[225].

Comprehensively defining biomarkers of translational value adopting a multi-omics and phenotypic approach alongside complementary image analysis studies harnessing modern spatial biology techniques that combine artificial intelligence and digital pathology could yield high-quality spatially resolved molecular profiles of MBC, improving our understanding of this rare cancer.

References:

We cited 239 references in this manuscript, including the 197 studies that met the inclusion criteria of the systematic review. The first 100 references are listed below with the rest in the Appendix (Page 130).

1. Fox S, Speirs V, Shaaban AM. Male breast cancer: an update. *Virchows Arch* 2022; **480**(1): 85-93.
2. Xu S, Liu Y, Zhang T, et al. The Global, Regional, and National Burden and Trends of Breast Cancer From 1990 to 2019: Results From the Global Burden of Disease Study 2019. *Front Oncol* 2021; **11**: 689562.
3. Ferzoco RM, Ruddy KJ. The Epidemiology of Male Breast Cancer. *Curr Oncol Rep* 2016; **1**(18): 1-6.
4. Gucalp A, Traina TA, Eisner JR, et al. Male breast cancer: a disease distinct from female breast cancer. *Breast Cancer Res Treat* 2019; **173**(1): 37-48.
5. Darkeh MHSE, Azavedo E. Male breast cancer clinical features, risk factors, and current diagnostic and therapeutic approaches. *Int J Clin Med* 2014; **5**(17): 1068-86.
6. Humphries MP, Rajan SS, Honarpisheh H, et al. Characterisation of male breast cancer: a descriptive biomarker study from a large patient series. *Sci Rep* 2017; **7**(1): 1-9.
7. Cardoso F, Bartlett JMS, Slaets L, et al. Characterization of male breast cancer: Results of the EORTC 10085/TBCRC/BIG/NABCG International Male Breast Cancer Program. *Ann Oncol* 2018; **29**(2): 405-17.
8. Cruz-Tapias P, Rubiano W, Rondón-Lagos M, Villegas VE, Rangel N. Intrinsic Subtypes and Androgen Receptor Gene Expression in Primary Breast Cancer. A Meta-Analysis. *Biology* 2021; **10**(9): 834.
9. Rozeboom B, Dey N, De P. ER+ metastatic breast cancer: past, present, and a prescription for an apoptosis-targeted future. *Am J Cancer Res* 2019; **9**(12): 2821-31.
10. Johansson I, Lauss M, Holm K, et al. Genome methylation patterns in male breast cancer - Identification of an epitype with hypermethylation of polycomb target genes. *Mol Oncol* 2015; **9**(8): 1565-79.
11. Zelli V, Silvestri V, Valentini V, et al. Transcriptome of Male Breast Cancer Matched with Germline Profiling Reveals Novel Molecular Subtypes with Possible Clinical Relevance. *Cancers* 2021; **13**(18): 4515.
12. Lacle MM, Kornegoor R, Moelans CB, et al. Analysis of copy number changes on chromosome 16q in male breast cancer by multiplex ligation-dependent probe amplification. *Mod Pathol* 2013; **26**(11): 1461-7.

13. Johansson I, Nilsson C, Berglund P, et al. Gene expression profiling of primary male breast cancers reveals two unique subgroups and identifies N-acetyltransferase-1 (NAT1) as a novel prognostic biomarker. *Breast Cancer Res* 2012; **14**(1): 1-5.
14. Johansson I, Nilsson C, Berglund P, et al. High-resolution genomic profiling of male breast cancer reveals differences hidden behind the similarities with female breast cancer. *Breast Cancer Res Treat* 2011; **129**(3): 747-60.
15. Kornegoor R, Moelans CB, Verschuur-Maes AHJ, et al. Promoter hypermethylation in male breast cancer: Analysis by multiplex ligation-dependent probe amplification. *Breast Cancer Res* 2012; **14**(4): 1-9.
16. Barnes DR, Silvestri V, Leslie G, et al. Breast and Prostate Cancer Risks for Male BRCA1 and BRCA2 Pathogenic Variant Carriers Using Polygenic Risk Scores. *J Natl Cancer Inst* 2022; **114**(1): 109-122.
17. Fentiman IS. Male breast cancer is not congruent with the female disease. *Crit Rev Oncol Hematol* 2016; **100**(101): 119-24.
18. Rizzolo P, Silvestri V, Tommasi S, et al. Male breast cancer: genetics, epigenetics, and ethical aspects. *Ann Oncol* 2013; **24**(Supplement 8): viii75-viii82
19. Fentiman I. Male breast cancer: a review. *Ecancermedicalscience* 2009; **3**: 140.
20. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *Syst Rev* 2021; **10**(1): 1-11.
21. Moola S, Munn Z, Tufanaru C, et al. Chapter 7: Systematic reviews of etiology and risk. In: Aromataris E, Munn Z, eds. *JBI Manual for Evidence Synthesis*. JBI, 2020.
22. Rizzolo P, Silvestri V, Valentini V, et al. Evaluation of CYP17A1 and CYP1B1 polymorphisms in male breast cancer risk. *Endocr Connect* 2019; **8**(8): 1224-9.
23. Di Oto E, Biserni GB, Varga Z, et al. X chromosome gain is related to increased androgen receptor expression in male breast cancer. *Virchows Arch* 2018; **473**(2): 155-63.
24. Karhu R, Laurila E, Kallioniemi A, Syrjäkoski K. Large genomic BRCA2 rearrangements and male breast cancer. *Cancer Detect Prev* 2006; **30**(6): 530-4.
25. Rudlowski C, Friedrichs N, Faridi A, et al. HER-2/neu gene amplification and protein expression in primary male breast cancer. *Breast Cancer Res Treat* 2004; **84**(3): 215-23.
26. Wick MR, Sayadi H, Ritter JH, Hill DA, Reddy VB, Gattuso P. Low-stage carcinoma of the male breast. A histologic, immunohistochemical, and flow cytometric comparison with localized female breast carcinoma. *Am J Clin Pathol* 1999; **111**(1): 59-69.
27. Prechtel D, Werenskiold AK, Prechtel K, Keller G, Hofler H. Frequent loss of heterozygosity at chromosome 13q12-13 with BRCA2 markers in sporadic male breast cancer. *Diagn Mol Pathol* 1998; **7**(1): 57-62.

28. Wingren S, van den Heuvel A, Gentile M, Olsen K, Hatschek T, Soderkvist P. Frequent allelic losses on chromosome 13q in human male breast carcinomas. *Eur J Cancer* 1997; **33**(14): 2393-6.
29. Chuaqui RF, Sanz-Ortega J, Vocke C, et al. Loss of Heterozygosity on the Short Arm of Chromosome 8 in Male Breast Carcinomas. *Cancer Res* 1995; **55**(21): 4995-8.
30. Biserni GB, Di Oto E, Moskovszky LE, Foschini MP, Varga Z. Preferential expression of NY-BR-1 and GATA-3 in male breast cancer. *J Cancer Res Clin Oncol* 2018; **144**(2): 199-204.
31. Kornegoor R, van Diest PJ, Buerger H, Korsching E. Tracing differences between male and female breast cancer: both diseases own a different biology. *Histopathology* 2015; **67**(6): 888-97.
32. Verschuur-Maes AH, Kornegoor R, de Bruin PC, Oudejans JJ, van Diest PJ. Do columnar cell lesions exist in the male breast? *Histopathology* 2014; **64**(6): 818-25.
33. Sanchez-Munoz A, Roman-Jobacho A, Perez-Villa L, et al. Male Breast Cancer: Immunohistochemical Subtypes and Clinical Outcome Characterization. *Oncology* 2012; **83**(4): 228-33.
34. Kraus TS, Cohen C, Siddiqui MT. Prostate-specific antigen and hormone receptor expression in male and female breast carcinoma. *Diagn Pathol* 2010; **5**(1): 1-6.
35. Shah P, Robbani I, Shah O. Clinicopathological study of male breast carcinoma: 24 Years of experience. *Ann Saudi Med* 2009; **29**(4): 288-93.
36. Jamal S, Mushtaq H, Mubarik A, Malik TM. Estrogen Receptor, Progesterone Receptor, HER2/neu, P53 and Ki-67 status of Male Breast Carcinomas in Pakistan. *Asian Pac J Cancer Prev* 2009; **10**(6): 1067-70.
37. Ge Y, Sneige N, Eltorky MA, et al. Immunohistochemical characterization of subtypes of male breast carcinoma. *Breast Cancer Res* 2009; **11**(3): 1-8.
38. Milias S, Kalekou H, Bobos M, et al. Immunohistochemical investigation of CD34 antigen in male breast carcinoma. *Clin Exp Med* 2007; **7**(3): 122-6.
39. Ciocca V, Bombonati A, Gatalica Z, et al. Cytokeratin profiles of male breast cancers. *Histopathology* 2006; **49**(4): 365-70.
40. Frangou EM, Lawson J, Kanthan R. Angiogenesis in male breast cancer. *World J Surg Oncol* 2005; **3**(1): 1-4.
41. Murphy CE, Carder PJ, Lansdown MRJ, Speirs V. Steroid hormone receptor expression in male breast cancer. *Eur J Surg Oncol* 2006; **32**(1): 44-7.
42. Kidwai N, Gong Y, Sun X, et al. Expression of androgen receptor and prostate-specific antigen in male breast carcinoma. *Breast Cancer Res* 2003; **6**(1): 1-6.
43. Blin N, Kardaś I, Welter C, et al. Expression of the C- erbB2 proto-oncogenein male breast carcinoma: Lack of prognostic significance. *Oncology* 1993; **50**(6): 408-11.

631 44. Kardaś I, Seitz G, Limon J, et al. Retrospective analysis of prognostic significance of
632 the estrogen-inducible pS2 gene in male breast carcinoma. *Cancer* 1993; **72**(5): 1652-6.
633 45. Alm P, Alumets J, Bak-Jensen E, Olsson H. Neuroendocrine differentiation in male
634 breast carcinomas. *Apmis* 1992; **100**(7-12): 720-6.
635 46. Li X, Wang J, Fu J, Xie X, Xie X. Prognostic significance of preoperative serum
636 inflammation markers in patients with male breast cancer. *Transl Cancer Res* 2021; **10**(9):
637 4002-8.
638 47. Hartkopf AD, Taran FA, Walter CB, et al. Detection and prevalence of disseminated
639 tumor cells from the bone marrow of early stage male breast cancer patients. *Breast Cancer*
640 *Res Treat* 2015; **152**(1): 51-5.
641 48. Bloom KJ, Govil H, Gattuso P, Reddy V, Francescatti D. Status of HER-2 in male and
642 female breast carcinoma. *Am J Surg* 2001; **182**(4): 389-92.
643 49. Szwiec M, Tomiczek-Szwiec J, Kluźniak W, et al. Genetic predisposition to male
644 breast cancer in Poland. *BMC Cancer* 2021; **21**(1): 1-8.
645 50. Ben Kridis-Rejeb W, Ben Ayed-Guerfali D, Ammous-Boukhris N, et al. Identification
646 of novel candidate genes by exome sequencing in Tunisian familial male breast cancer
647 patients. *Mol Biol Rep* 2020; **47**(9): 6507-16.
648 51. Vietri MT, Caliendo G, D'Elia G, et al. BRCA and PALB2 mutations in a cohort of
649 male breast cancer with one bilateral case. *Eur J Med Gen* 2020; **63**(6): 103883.
650 52. Tedaldi G, Tebaldi M, Zampiga V, et al. Male Breast Cancer: Results of the
651 Application of Multigene Panel Testing to an Italian Cohort of Patients. *Diagnostics* 2020;
652 **10**(5): 269.
653 53. Scarpitta R, Zanna I, Aretini P, et al. Germline investigation in male breast cancer of
654 DNA repair genes by next-generation sequencing. *Breast Cancer Res Treat* 2019; **178**(3):
655 557-64.
656 54. Rizzolo P, Zelli V, Silvestri V, et al. Insight into genetic susceptibility to male breast
657 cancer by multigene panel testing: Results from a multicenter study in Italy. *Int J Cancer*
658 2019; **145**(2): 390-400.
659 55. Silvestri V, Rizzolo P, Zelli V, et al. A possible role of FANCM mutations in male
660 breast cancer susceptibility: Results from a multicenter study in Italy. *Breast* 2018; **38**: 92-7.
661 56. Fostira F, Saloustros E, Apostolou P, et al. Germline deleterious mutations in genes
662 other than BRCA2 are infrequent in male breast cancer. *Breast Cancer Res Treat* 2018;
663 **169**(1): 105-13.
664 57. Schayek H, Korach H, Laitman Y, Bernstein-Molho R, Friedman E. Mutational
665 analysis of candidate genes in Israeli male breast cancer cases. *Breast Cancer Res Treat*
666 2018; **170**(2): 399-404.

58. Hallamies S, Pelttari LM, Poikonen-Saksela P, et al. CHEK2 c.1100delC mutation is associated with an increased risk for male breast cancer in Finnish patient population. *BMC Cancer* 2017; **17**(1): 1-5.
59. Deb S, Jene N, investigators K, Fox SB. Genotypic and phenotypic analysis of familial male breast cancer shows under representation of the HER2 and basal subtypes in BRCA-associated carcinomas. *BMC Cancer* 2012; **12**(1): 1-3.
60. Silvestri V, Rizzolo P, Falchetti M, et al. Mutation analysis of BRIP1 in male breast cancer cases: a population-based study in Central Italy. *Breast Cancer Res Treat* 2011; **126**(2): 539-43.
61. Ding YC, Steele L, Kuan CJ, Greilac S, Neuhausen SL. Mutations in BRCA2 and PALB2 in male breast cancer cases from the United States. *Breast Cancer Res Treat* 2011; **126**(3): 771-8.
62. Besic N, Cernivc B, de Greve J, et al. BRCA2 gene mutations in Slovenian male breast cancer patients. *Genet Test* 2008; **12**(2): 203-9.
63. Wasielewski M, Den Bakker MA, Van Den Ouweland A, et al. CHEK2 1100delC and male breast cancer in the Netherlands. *Breast Cancer Res Treat* 2009; **116**(2): 397-400.
64. Chodick G, Struewing JP, Ron E, Rutter JL, Iscovich J. Similar prevalence of founder BRCA1 and BRCA2 mutations among Ashkenazi and non-Ashkenazi men with breast cancer: Evidence from 261 cases in Israel, 1976-1999. *Eur J Med Gen* 2008; **51**(2): 141-7.
65. Falchetti M, Lupi R, Rizzolo P, et al. BRCA1/BRCA2 rearrangements and CHEK2 common mutations are infrequent in Italian male breast cancer cases. *Breast Cancer Res Treat* 2008; **110**(1): 161-7.
66. Tchou J, Ward MR, Volpe P, et al. Large genomic rearrangement in BRCA1 and BRCA2 and clinical characteristics of men with breast cancer in the United States. *Clin Breast Cancer* 2007; **7**(8): 627-33.
67. Syrjäkoski K, Fredriksson H, Ikonen T, et al. Hemochromatosis gene mutations among Finnish male breast and prostate cancer patients. *Int J Cancer* 2006; **118**(2): 518-20.
68. Miolo G, Della Puppa L, Santarosa M, et al. Phenotypic features and genetic characterization of male breast cancer families: identification of two recurrent BRCA2 mutations in north-east of Italy. *BMC Cancer* 2006; **6**(1): 1-9.
69. Syrjakoski K, Kuukasjarvi T, Waltering K, et al. BRCA2 mutations in 154 finnish male breast cancer patients. *Neoplasia* 2004; **6**(5): 541-5.
70. Tournier I, Brigitte BDP, Sobol H, et al. Significant contribution of germline BRCA2 rearrangements in male breast cancer families. *Cancer Res* 2004; **64**(22): 8143-7.
71. Ottini L, Masala G, D'Amico C, et al. BRCA1 and BRCA2 mutation status and tumor characteristics in male breast cancer: A population-based study in Italy. *Cancer Res* 2003; **63**(2): 342-7.

72. Gudmundsdottir K, Thorlacius S, Jonasson JG, Sigfusson BF, Tryggvadottir L, Eyfjord JE. CYP17 promoter polymorphism and breast cancer risk in males and females in relation to BRCA2 status. *Br J Cancer* 2003; **88**(6): 933-6.
73. Evans DG, Bulman M, Young K, Gokhale D, Lalloo F. High detection rate for BRCA2 mutations in male breast cancer families from North West England. *Fam Cancer* 2001; **1**(3-4): 131-3.
74. Basham VM, Lipscombe JM, Ward JM, et al. BRCA1 and BRCA2 mutations in a population-based study of male breast cancer. *Breast Cancer Res* 2002; **4**(1): 1-5.
75. Wolpert N, Warner E, Seminsky MF, Futreal A, Narod SA. Prevalence of BRCA1 and BRCA2 mutations in male breast cancer patients in Canada. *Clin Breast Cancer* 2000; **1**(1): 57-63.
76. Díez O, Cortés J, Domènech M, et al. BRCA2 germ-line mutations in Spanish male breast cancer patients. *Ann Oncol* 2000; **11**(1): 81-4.
77. Csokay B, Udvarhelyi N, Sulyok Z, et al. High frequency of germ-line BRCA2 mutations among Hungarian male breast cancer patients without family history. *Cancer Res* 1999; **59**(5): 995-8.
78. Haraldsson K, Loman N, Zhang QX, Johannsson O, Olsson H, Borg A. BRCA2 germ-line mutations are frequent in male breast cancer patients without a family history of the disease. *Cancer Res* 1998; **58**(7): 1367-71.
79. Mavraki E, Gray IC, Bishop DT, Spurr NK. Germline BRCA2 mutations in men with breast cancer. *Br J Cancer* 1997; **76**(11): 1428-31.
80. Friedman LS, Gayther SA, Kurosaki T, et al. Mutation analysis of BRCA1 and BRCA2 in a male breast cancer population. *Am J Hum Genet* 1997; **60**(2): 313-9.
81. Thorlacius S, Sigurdsson S, Bjarnadottir H, et al. Study of a single BRCA2 mutation with high carrier frequency in a small population. *Am J Hum Genet* 1997; **60**(5): 1079-84.
82. Anelli A, Anelli TFM, Youngson B, Rosen PP, Borgen PI. Mutations of the p53 gene in male breast cancer. *Cancer* 1995; **75**(9): 2233-8.
83. Deb S, Gorringe KL, Pang JMB, et al. BRCA2 carriers with male breast cancer show elevated tumour methylation. *BMC Cancer* 2017; **17**(1): 1-11.
84. Kornegoor R, Moelans CB, Verschuur-Maes AHJ, et al. Oncogene amplification in male breast cancer: analysis by multiplex ligation-dependent probe amplification. *Breast Cancer Res Treat* 2012; **135**(1): 49-58.
85. Zanna I, Silvestri V, Palli D, et al. Smoking and FGFR2 rs2981582 variant independently modulate male breast cancer survival: A population-based study in Tuscany, Italy. *Breast* 2018; **40**: 85-91.

86. Chen L, Weng YM, Hu MX, Peng M, Song QB. Effects of HER2 status on the prognosis of male breast cancer: A population-based study. *Onco Targets Ther* 2019; **12**: 7251-60.
87. André S, Pereira T, Silva F, et al. Male breast cancer: Specific biological characteristics and survival in a Portuguese cohort. *Mol Clin Oncol* 2019; **10**(6): 644-54.
88. Moelans CB, De Ligt J, Van Der Groep P, et al. The molecular genetic make-up of male breast cancer. *Endocr Relat Cancer* 2019; **26**(10): 779-94.
89. Gargiulo P, Pensabene M, Milano M, et al. Long-term survival and BRCA status in male breast cancer: A retrospective single-center analysis. *BMC Cancer* 2016; **16**(1): 1-11.
90. Piscuoglio S, Ng CK, Murray MP, et al. The Genomic Landscape of Male Breast Cancers. *Clin Cancer Res* 2016; **22**(16): 4045-56.
91. Abreu MH, Gomes M, Menezes F, et al. CYP2D6 4 polymorphism: A new marker of response to hormonotherapy in male breast cancer? *Breast* 2015; **24**(4): 481-6.
92. Johansson I, Ringner M, Hedenfalk I. The Landscape of Candidate Driver Genes Differs between Male and Female Breast Cancer. *PLoS One* 2013; **8**(10): e78299.
93. Kornegoor R, Verschuur-Maes AHJ, Buerger H, et al. Immunophenotyping of male breast cancer. *Histopathology* 2012; **61**(6): 1145-55.
94. Song YN, Geng JS, Liu T, et al. Long CAG Repeat Sequence and Protein Expression of Androgen Receptor Considered as Prognostic Indicators in Male Breast Carcinoma. *PLoS One* 2012; **7**(12): e52271.
95. Fonseca RR, Tomas AR, Andre S, Soares J. Evaluation of ERBB2 gene status and chromosome 17 anomalies in male breast cancer. *Am J Surg Pathol* 2006; **30**(10): 1292-8.
96. Kwiatkowska E, Teresiak M, Filas V, Karczewska A, Bręborowicz D, Mackiewicz A. BRCA2 mutations and androgen receptor expression as independent predictors of outcome of male breast cancer patients. *Clin Cancer Res* 2003; **9**(12): 4452-9.
97. Zhang S, Liu B, Zhou M, et al. The beneficial role of Asian-based RecurIndex test in the prognostic prediction in Chinese male breast cancer patients. *Sci Rep* 2021; **11**(1): 1-8.
98. Severson TM, Kim Y, Joosten SEP, et al. Characterizing steroid hormone receptor chromatin binding landscapes in male and female breast cancer. *Nat Commun* 2018; **9**(1): 1-12.
99. Massarweh SA, Sledge GW, Miller DP, McCullough D, Petkov VI, Shak S. Molecular Characterization and Mortality From Breast Cancer in Men. *J Clin Oncol* 2018; **36**(14): 1396-404.
100. André S, Pinto AE, Silva GL, Silva F, Serpa J, Félix A. Male Breast Cancer-Immunohistochemical Patterns and Clinical Relevance of FASN, ATF3, and Collagen IV. *Breast Cancer: Basic Clin Res* 2021; **15**: 1-14.

Figure legend

Figure 1

(A) MBC biomarkers that were investigated across multiple omics categories aligned to their associated survival outcomes if present; (B) MBC biomarkers that had associations with multiple hallmarks of cancer aligned to their associated survival outcomes if present.

Table 1: (A) common proteomic biomarkers in breast cancer, **(B)** other well-studied proteomic biomarkers in MBC and their effects on prognosis

Protein biomarkers	**Effects on prognosis**
(A) Common biomarkers	
ERα	***Positivity predicts:*** Improved OS* (frequency = 99.3%[7], 87.6%[104], and 32%[134]; all $p < 0.05$)[7,104,134]; improved DFS* (frequency = 99.3%; $p = 0.001$)[7]; improved DSS* (frequency = 93%; $p < 0.01$)[121] ***Positivity associated with:*** *L*ow Ki-67 index (frequency = 93.1%[87] and 91%[133]; both $p < 0.05$)[87,133]; PR positivity (frequency = 82%; $p = 0.01$)[202]; AR positivity (frequency = 91%; $p = 0.036$)[133]; Bcl-2 positivity (frequency = 82%; $p = 0.04$)[202]; pS2 positivity (frequency = 82%; $p = 0.04$)[202]; >60 years of age at diagnosis (frequency = 82%; $p = 0.03$)[202] ***More frequently expressed in:*** MBCs* compared to FBCs* in general (frequency = 100% vs 86%[136] and 82.3% vs 53.4%[185]; both $p < 0.05$)[136,185]; MBCs compared to post-menopausal FBCs* (frequency = 82.3% vs 48.9%; $p = 0.01$)[185] ***Other:*** Lower intensity of expression in MBCs* compared to FBCs* of age group 26-35 years ($p = 0.001$)[191]; higher median tumour levels in MBCs* compared to FBCs* ($p = 0.02$)[135]
PR	***Positivity predicts:*** Improved OS* (frequency = 81.9%[7]; 67.2%[104], and 80%[105]; all $p < 0.05$)[7,104,105]; improved DFS* (frequency = 81.9%; $p = 0.002$)[7]; improved DSS* (frequency = 77%; $p = 0.01$)[121]; reduced OS* ($p = 0.036$)[103]**; reduced DFS* ($p = 0.01$)[103]** ***Positivity associated with:*** Low Ki-67 index ($p < 0.001$); low pathological stage ($p = 0.029$); *BRCA2* mutation negativity ($p = 0.01$). Frequency = 75.2%[87] ***Other:*** Higher frequency of positivity in MBCs* compared to FBCs* (frequency = 91% vs 76%[136] and 77% vs 62%[202]; $p = 0.01$)[136,202]; lower intensity of expression in MBCs* compared to FBCs* of age group 26-35 years ($p = 0.001$)[191]; higher median tumour levels in MBCs* compared to FBCs* ($p = 0.04$)[135]
ERα/PR co-expression	***Positivity predicts:*** Improved OS* (frequency = 78.1%; $p = 0.0054$)[118]; improved DFS* ($p = 0.022$)[118] ***Positivity associated with:*** Low Ki-67 index (frequency = 78.1%; $p = 0.029$)[118]

HER2	***Positivity predicts:*** Reduced OS* (frequency = 8%[95], 13.5%[101], and 56%[129]; all p < 0.05)[95,101,129]; reduced OS* in ERα positive cases (p = 0.003)[6]; reduced DSS* (p = 0.0001)[101] ***Positivity associated with:*** Younger age of diagnosis (frequency = 13.5%; p < 0.001)[101]; large tumour size (frequency = 3%; p < 0.001)[188]; distant metastasis (frequency = 11%; p = 0.009)[87]; high Ki-67 index (frequency = 11%; p = 0.011)[87]; high anatomic stage (frequency = 11%; p = 0.015)[87]; high tumour grade (frequency = 3%[188] and 62.5%[198]; both p < 0.05)[188,198]
AR	***Positivity predicts:*** Improved OS* in general (frequency = 96.9%[7] and 62.5%[116]; both p < 0.05)[7,116]; improved DFS* in general (frequency = 96.9%[7]; both p < 0.05)[6,7]**; improved 5-year OS* in Luminal A MBCs* compared to Luminal A FBCs* (frequency = 64%; p = 0.01)[123]; reduced 5-year OS* in general (frequency = 82.7%[94], 55.8%[96], and 40.2%[117]; all p < 0.05)[94,96,117]; reduced DFS* in general (frequency = 55.8%; p = 0.002)[96]; reduced 5-year DFS* (frequency = 82.7%[94] and 40.2%[117]; both p < 0.05)[94,117] ***Positivity associated with:*** ERα positivity (frequency = 82.7%[94], 62.5%[116], and 34%[131]; all p < 0.05)[94,116,131,179]**; PR positivity (frequency = 82.7%; p = 0.024)[94]; older age at diagnosis (frequency = 38.5%; p = 0.05)[200]; low proliferative activity (frequency = 34%; p = 0.04)[131]; low tumour grade (p < 0.05)[179]**; poor clinical benefit (frequency = 40.2%; p = 0.025)[117]; node positivity (frequency = 40.2%; p = 0.032)[117]; node negativity in cases with <20% PR positivity (p = 0.007)[179]** ***Other:*** Higher frequency of positivity in MBCs* compared to FBCs* (frequency = 94% vs 63%; p < 0.0001)[123]
Ki-67/MIB1	***High Ki-67 / MIB-1 index predicts:*** Reduced OS* (frequency = 58.9%[87], 48%[129], 46.8%[131], and 48.2%[135]; all p < 0.05)[87,129,131,135]; reduced DFS* (frequency = 58.9%; p = 0.03)[87]; reduced PFS* (frequency = 38%; p = 0.012)[133] ***High Ki-67 / MIB-1 index associated with:*** High tumour grade (frequency = 58.9%[87] and 46.9%[118]; all p < 0.05)[87,118,186,196]**; high anatomic stage (frequency = 58.9%; p = 0.004)[87]; node positivity (frequency = 58.9%[87] and 19.4%[197]; both p < 0.01)[87,197]; positive family history (frequency = 58.9%; p = 0.002)[87]; *BRCA2* mutation positivity (frequency = 58.9%; p = 0.047)[87]; ERα/PR co-expression (both p < 0.05)[186,200]**
(B) Other biomarkers	**Effects on prognosis**
p53	***Positivity predicts:*** Reduced 10-year OS (frequency = 21.2%; p = 0.015)[119] ***Positivity associated with:*** ERα negativity (frequency = 13.6%; p = 0.002)[202]; PR negativity (frequency = 13.6%; p < 0.001)[202]; Bcl-2 negativity (frequency = 13.6%; p = 0.02)[202]; node metastases (frequency = 15%[93] and 16.7%[181]; both p < 0.05)[93,181]; tumour grade 3 (overexpression) (frequency = 15%; p = 0.049)[93]

	Other: Positivity[128,129,131] / overexpression[93] independently predicts reduced OS (frequency = 54%[128], 54%[129], 57.4%[131], and 15%[93]; all $p < 0.05$); negativity associated with Luminal A type (frequency = 78.8%[119] and 83.3%[181]; both $p < 0.05$)[119,181]; higher frequency of positivity in FBCs compared to MBCs (frequency = 18% vs 4%; $p < 0.001$)[160]
Bcl-2	***Positivity associated with:*** ERα positivity (frequency = 94%; $p = 0.04$)[189]; PR positivity (frequency = 56.6%; $p = 0.008$)[194]; node positivity (frequency = 66.7%[181] and 56.6%[194]; both $p < 0.05$)[181,194]; small tumour size (frequency = 73%; $p = 0.017$)[93] ***Negativity associated with***: Luminal B type ($p = 0.028$); tumour grade 3 ($p = 0.01$), frequency = 25%[93] ***Other:*** Higher frequency of positivity in MBCs* compared to FBCs* (frequency = 67% vs 48%; $p = 0.006$)[202]
Cyclin D1	***Positivity predicts:*** Improved PFS* (frequency = 58%; $p = 0.009$)[133]; improved DFS* (frequency = 83.7%; $p = 0.04$)[125]; improved DSS* ($p = 0.001$)[121]** ***Positivity associated with:*** Small tumour size (frequency = 77%[93] and 83.7%[125]; both $p < 0.05$)[93,125]; node negativity (frequency = 83.7%; $p = 0.04$)[125]; p53 positivity (frequency = 58%; $p < 0.001$)[133]; AR positivity (frequency = 58%; $p = 0.028$)[133]
Hypoxic biomarkers **HIF1-α**	***Positivity predicts:*** Reduced DSS* in sporadic MBCs* but not familial MBCs* (frequency = 59% vs 15.5%; $p = 0.006$)[141]; overexpression independently predicts reduced DSS* (frequency = 27%; $p < 0.05$)[124]; perinecrotic staining predicts reduced OS* (frequency = 22.4%; $p = 0.014$)[124†]; diffuse staining in >5% tumour cells associated with high histological grade ($p < 0.001$) and high mitotic count ($p = 0.038$; frequency = 34.4%)[124] ***Positivity associated with***: Invasive carcinoma of no special type ($p = 0.005$); basal cell intrinsic phenotype ($p = 0.02$; frequency = 25.1%)[141] ***Overexpression associated with:*** High tumour grade (frequency = 27%[124] and 36.2%[180]; both $p < 0.05$)[124,180]; high mitotic activity (frequency = 36.2%; $p = 0.013$)[180]; HER2 amplification (frequency = 27%; $p = 0.005$)[124]; Glut-1 overexpression (frequency = 27%; $p < 0.001$)[124]; CA-9 overexpression (frequency = 27%; $p = 0.034$)[124] ***Other:*** High similarity of expression between invasive carcinoma and adjacent DCIS* (frequency = 36.2% vs 37.9%; $p < 0.001$)[180]; higher frequency of Glut-1/CA-9 overexpression with HIF1-α perinecrotic staining compared to diffuse staining in DCIS* (both pure and adjacent) (frequency = 60% vs 100%; $p = 0.012$)[180];

CA-9	***Positive expression predicts:*** Reduced DSS* (frequency = 8%; p = 0.002)[141] ***Other:*** High similarity of expression between invasive carcinoma and adjacent DCIS* (frequency = 37.9% vs 24.1%; p < 0.001)[180]
HIF1-α and/or CA-9 expression	***Expression of either marker predicts:*** Reduced DSS* (frequency = 25.1% and 8% for HIF1-α and CA-9 respectively; p = 0.008)[141]
Glut-1	***Overexpression associated with:*** High mitotic count (p = 0.014); high tumour grade (p = 0.038; frequency = 62.1% for invasive carcinoma[180] ***Other:*** High similarity of expression between invasive carcinoma and adjacent DCIS* (frequency = 75.8% vs 62.1%; p < 0.001)[180]
p21	***Positivity predicts:*** Reduced DFS* (frequency = 41.3%; p = 0.04)[125] ***Positivity associated with***: HER2 negativity (frequency = 70.3%; p = 0.05)[196]; high mitotic activity (frequency = 48%; p < 0.001)[93]; tumour grade 3 (frequency = 48%; p = 0.002)[93] ; Luminal B type (frequency = 48%; p = 0.026)[93] ***Other:*** Higher frequency of positivity in MBCs* compared to FBCs* (frequency = 96% vs 58%[160] and 70.3% vs 29%[196]; both p < 0.01)[160,196]
p27	***Negativity associated with:*** Lymph node metastases (frequency = 81.2%[125] and 64%[197]; both p < 0.05)[125,197] ***Overexpression associated with:*** AR positivity (frequency = 96.2%; p = 0.049)[196] ***Other:*** Higher frequency of positivity in MBCs* compared to FBCs* (frequency = 96.2% vs 39.3%; p = 0.00)[196]
EGFR	***Overexpression associated with***: HER2 amplification (frequency = 12%; p = 0.04)[190] ***Positivity associated with:*** ERα and PR negativity (frequency = 11.4%; both p = 0.04)[188]; high MIB-1 index (frequency = 9.4%; p = 0.0181)[118]
c-Myc	***Positivity predicts:*** Reduced OS* (frequency = 82%; p = 0.01)[129]

	Other: Overexpression predicts improved DFS* (frequency = 90%; p = 0.04)[125] and is associated with node negativity (frequency = 90%; p = 0.006)[125]

**MBC: Male Breast Cancer; FBC: Female Breast Cancer; OS: Overall Survival; DFS: Disease Free Survival; DSS: Disease Specific Survival; PFS: Progression Free Survival; DCIS: Ductal Carcinoma In-Situ*

***frequency unavailable from all/some source article(s)*

†*Perinecrotic staining: Staining surrounding a necrotic area*

Table 2: Ten most studied genetic/transcriptomic biomarkers in MBC and their effects on prognosis

Biomarker	Effects on prognosis
BRCA2	***Mutated status predicts:*** Reduced OS* in general (frequency = 10.8%[85] and 29.5%[87]; both p < 0.05)[85,87]; reduced 5-year OS* (frequency = 27.9%; p = 0.003)[96]; reduced DSS* in general (frequency = 29.5%; p = 0.003)[87]; reduced 5-year DSS* (frequency = 27.9%; p = 0.006)[96] ***Mutated status associated with:*** ERα negativity (frequency = 9.3%; p = 0.05)[173]; PR negativity (frequency = 29.5%[87], 12.2%[170] and 9.3%[173]; all p < 0.05)[87,170,173;] HER2 positivity/enriched subtype (frequency = 12.2%[170] and 9.3%[173]; both p < 0.05)[170,173]; Luminal B type (frequency = 12.2%; p = 0.016)[170]; advanced tumour grade[164,173]/ tumour grade 3[61,170] (frequency = 89.4%[164†], 9.3%[173], 15.6%[61], and 12.2%[170]; all p < 0.05); higher frequency of tumour grade 3 in patients <50 years of age (frequency = 89.4%; p = 0.005)[164†]; node positivity (frequency = 15.6%; p < 0.02)[61]; contralaterality (frequency = 12.2%; p = 0.01)[170]; bilaterality (frequency = 29.5%; p = 0.008)[87]; high Ki-67 index (frequency = 29.5%; p = 0.047)[87]; higher frequency of genetic aberrations in *BRCA2*-mutated MBCs compared to *BRCA2*-wt MBCs (p < 0.05)[175]**; family history of breast/ovarian cancer or personal history of cancer (frequency = 12.2%[170]; all p < 0.05)[68,170]**; amplification of *CCNE2, ASAP1, CSMD3, UBR5, DNAH11, RRM2B, FZD6, RUNX1T1* and *SGK3* (frequency = 11%; all p < 0.05)[88]; decreased copy number aberration load on chr 8p (frequency = 11%; p = 0.004)[88] ***Other:*** Higher frequency of mutations in MBCs* compared to FBCs* (frequency = 41.7% vs 8.3%; p = 0.0008)[59]; higher tumour grade in *BRCA2*-mutated MBCs* compared to SEER* MBCs* (p = 4.52e-12)[164]; higher disease stage in *BRCA2*-mutated MBCs* compared to *BRCA2*-mutated FBCs* (p = 2.14e-5)[164]; increased disease risk in men <60 years (OR* = 5.63; frequency = 29.4%; p < 0.05)[149]
HER2	***Amplified status predicts****:* Reduced OS* in general[86,95] – also predicted by copy number gain[84] (frequency = 13.3%[86], 8%[95], and 4%[84]; all p < 0.05); reduced 4-year OS* (frequency = 13.3%; p = 0.005)[86]; reduced OS* in patients with tumour size of 2-4 cm (frequency = 13.3%; p = 0.02)[86]; reduced OS* in patients with distant metastasis (frequency = 13.3%; p = 0.023)[86]; reduced OS* in patients who have undergone radiation therapy (frequency = 13.3%; p = 0.041)[86] ***Amplified status associated with****:* High mean mitotic activity (frequency = 3%; p < 0.001)[93]; poor degree of differentiation[86] / histological grade 3[93] (frequency = 13.3%[86] and 3%[93]; both p < 0.05); distant metastasis (frequency = 13.3%; p = 0.002)[86]; regional lymph node metastasis (frequency = 13.3%; p = 0.004)[86]; younger age of diagnosis (frequency = 13.3%; p < 0.001)[86]; large tumour size (frequency = 13.3%; p < 0.001)[86]; advanced disease stage (frequency = 13.3%; p < 0.001)[86]; surgery and chemotherapeutic treatment (frequency = 13.3%; p < 0.001)[86] ***Other:*** Downregulated in MBCs* compared to FBCs* (p < 0.01)[171]**

CCND1	***Amplified status associated with:*** ERα positivity (frequency = 63%; p < 0.0001)[174]; HER2 positivity (frequency = 16%; p = 0.0005)[165]; high MIB-1 index (frequency = 16%; p = 0.04)[165] ***Amplified status predicts***: Reduced OS* (frequency = 46%; p = 0.022)[84] ***Other:*** Higher copy number ratio and amplification frequency in high grade invasive carcinoma compared to low/intermediate grade invasive carcinoma (all p = 0.005)[162]**
PALB2	***Associations with MBC risk:*** Pathogenic variants associated with MBC risk (control dataset specific results; frequency = 1.2%)[54]; EVS* dataset: OR = 17.30 (p < 0.0001); ExAc* dataset: OR = 11.20 (p < 0.0001); gnomAD* dataset: OR = 9.63 (p < 0.0001) ***Other:*** Copy number gain (exon 6) associated with node negativity (p = 0.021)[12]**; Mutated status associated with bilaterality (frequency = 2.4%; p = 0.004)[49]; Higher frequency of mutations in MBC* compared to unmatched female normal breast tissue (frequency = 2.4%; p < 0.001)[49]
PIK3CA	***Mutated status associated with:*** *BRCA2* mutation negativity (frequency = 10.5%; p = 0.03)[169]; node positivity (frequency = 36.1%; p = 0.006)[88]; advanced tumour grade (frequency = 36.1%; p = 0.013)[88]; high mitotic index (frequency = 36.1%; p = 0.014)[88]; absence of both nuclear and cytoplasmic expression of p4E-BP1 (frequency = 10.5%; both p < 0.05)[169]; pS6 upregulation (frequency = 10.5%; p = 0.024)[169] ***Less frequently mutated in:*** ERα positive/HER2 negative MBCs* compared to matched FBCs* (frequency = 18% vs 42%; p = 0.0005)[90]; ERα positive/HER2 negative MBCs* compared to matched post-menopausal FBCs* (frequency = 18% vs 42%; p = 0.0014)[90]
GATA3	***Mutated status***: predicts reduced DFS* (frequency = 15%; p = 0.038)[90]; associated with Luminal B type (frequency = 15%; p = 0.0482)[90] ***Other:*** Upregulation associated with AR positivity (p = 0.0347)[171]**
EGFR	***Amplification associated with:*** ERα negativity (p = 0.01); HER2 positivity (p = 0.03); stage IV disease (p = 0.01). Amplification frequency = 6.8%[165] ***Other:*** Copy number gain associated with high grade invasive carcinoma (frequency = 62%; p = 0.047)[162]††
EMSY	***Amplification predicts:*** Reduced OS* (p = 0.04)[84]** ***Amplification associated with:*** *BRCA1/2* mutation positivity (frequency = 34.7%; p = 0.03)[163]

miR-125b	***High expression:*** associated with small tumour size (p = 0.03)[209]** ***Downregulated:*** MBCs* compared to FBCs* (p < 0.01); MBCs* compared to gynaecomastia (p < 0.01)[177]
rs3803662 (*TOX3;* risk biomarker)	**Associated with MBC* risk:** OR* = 1.48 (p = 4e-6)[145]**; OR* = 1.59 (frequency = 34.7%, 47.3%, and 18% for CC, CT, and TT genotypes, respectively; p = 0.0001)[144]

**MBC: Male Breast Cancer; FBC: Female Breast Cancer; OS: Overall Survival; DFS: Disease Free Survival; DSS: Disease Specific Survival; SEER: Surveillance Epidemiology and End Results; EVS: Exome Variant Server; ExAC: Exome Aggregation Consortium; gnomAD: Genome Aggregation Database*

***Breakdown for gene-specific alteration unavailable from all or some source articles*

†Cohort selected for BRCA1/2 mutations

††Frequency of CNV in pure ductal carcinoma in-situ (DCIS): 6% (CCND1 amplification), 6% (EGFR gain) and in DCIS adjacent to invasive carcinoma (DCIS-AIC): 16% (CCND1 amplification), 2% (EGFR gain)

Table 3: Ten most studied epigenetic biomarkers in MBC and their effects on prognosis

Biomarker	Effects on prognosis
ESR1	***Promoter hypermethylation:*** Associated with high tumour grade (p = 0.037); high mean mitotic count (p = 0.001), frequency = 8%[15] ***Other:*** Promoter hypermethylation less frequent in MBC* compared FBC* (frequency = 8%; p = 0.005)[15]; higher methylation in tumours compared to peripheral blood (p < 0.0001)[156]**; lower absolute methylation % in male DCIS-AIC* compared to female DCIS-AIC* (frequency of hypermethylated cases† in male DCIS-AIC = 5%; p < 0.002)[155]
GSTP1	***Promoter hypermethylation:*** Associated with high tumour grade (frequency = 44%; p = 0.001)[15]; high mean mitotic count (frequency = 44%; p = 0.002)[15]; *BRCA2* mutation positivity (frequency = 82%; p = 0.02)[83] ***Other:*** High absolute methylation % associated with high grade invasive carcinoma (frequency = 41%; p = 0.047)[155]
RARB	***Promoter hypermethylation:*** Associated with ERα negativity (frequency = 8%; p = 0.04)[157]; PR positivity (frequency = 8%; p = 0.03)[157]; large tumour size (frequency = 30%; p = 0.01)[83]; presence of Paget's disease (frequency = 30%; p = 0.01)[83]; *BRCA2* mutation positivity (frequency = 30%; p = 0.02)[83]; less frequent in MBC* compared FBC* (frequency = 5% vs 20%; p = 0.026)[15]
RASSF1/RASSF1A	***Promoter hypermethylation:*** Associated with ERα negativity (frequency = 76%; p = 0.0001)[157] ; PR positivity (frequency = 76%; p = 0.00)[157] ; HER2 positivity (frequency = 79.1%; p = 0.01)[156]; presence of DCIS* (frequency = 68%; p = 0.02)[83]; *BRCA1/2* mutation positivity (frequency = 79.1%; p = 0.008)[156]; tumour grade G3 (frequency = 79.1%; p = 0.008)[156]; more frequent in MBC* compared to FBC* (frequency = 76% vs 28%; p = 0.0001)[157] ***Other:*** Higher methylation levels in tumours compared to peripheral blood (p < 0.0001)[156]
AR	***Promoter hypermethylation:*** Associated with *BRCA1/2* mutation negativity (frequency = 94%; p = 0.016)[156] ***Other:*** CpG hypomethylation in MBC* cases compared to gynaecomastia cases (p < 0.05)[154.] Higher methylation in tumours compared to male normal breast tissue (p = 0.0009); tumours compared to lymph nodes (p = 0.003); tumours compared to peripheral blood (p = 0.0006). Frequency = 94%[156]
ATM	***Promoter hypermethylation:*** Less frequent in MBC* compared FBC* (frequency = 1% vs 15%; p = 0.017)[15] ***Other:*** High absolute methylation % associated with high grade invasive carcinoma (p = 0.036)[155]††

BRCA2	***Promoter hypermethylation:*** Less frequent in MBC* compared FBC* (frequency = 17% vs 60%; p < 0.001)[15] ***Other:*** Lower absolute methylation % in male DCIS-AIC* compared to female DCIS-AIC* (p < 0.02)[155]
MGMT	***Promoter hypermethylation:*** Associated with larger mean tumour size than tumours without *MGMT* hypermethylation (frequency = 7%; p = 0.002)[15]; higher frequency in pure invasive carcinoma compared to DCIS-AIC* (frequency = 25% vs 9%; p = 0.039)[155]
VHL	***Promoter hypermethylation:*** Less frequent in MBC* compared to FBC* (frequency = 2% vs 15%; p = 0.025)[15] ***Other:*** Lower absolute methylation % in male DCIS-AIC* compared to female DCIS-AIC* (p < 0.002)[155††]
TWIST1	***Promoter hypermethylation predicts:*** Reduced DSS* in *BRCA2* mutation positive MBC patients (p = 0.001); reduced DSS* in all MBC patients (p = 0.01). Frequency = 37%[83]

**MBC: Male Breast Cancer; FBC: Female Breast Cancer; DSS: Disease Specific Survival; DCIS: Ductal Carcinoma In-Situ; DCIS-AIC: Ductal Carcinoma In-situ Adjacent to Invasive Carcinoma*

***Frequency unavailable from source article*

†Frequency of ESR1 hypermethylated cases in male pure-DCIS = 6% and invasive carcinoma = 9%; frequency of BRCA2 hypermethylated cases in male pure-DCIS = 11% and invasive carcinoma = 2%

††Promoter hypermethylation was not present in the MBC cohort. However, higher absolute methylation % of ATM was observed in high grade tumours compared to low/intermediate grade tumours. Similarly, lower absolute methylation % of VHL was observed in male DCIS-AIC compared to female DCIS-AIC

Table 4: Ten most studied morphological features in MBC and their effects on prognosis

Morphological feature	Effects on prognosis
TIL* density	***High density of TILs*:*** Predicts improved OS* (p = 0.011) and RFS* (p =0.02, frequency = 14.3%)[137]; association with node positivity (frequency = 27.8%; p = 0.025)[186] ***Other:*** Higher density of TILs* in HER2 positive MBCs* compared to Luminal HER2 negative MBCs* (overall frequency of high TIL* density = 14.3%; p = 0.015)[137††]
Fibrotic focus	***Presence of fibrotic foci:*** Predicts reduced OS* (p = 0.004) and RFS* (p < 0.001) at a frequency of 32.2%)[137]; reduced overall survival when foci of >8 mm[†] (p = 0.035)[124] and associated with (frequency = 25%)[124]; high tumour grade (p = 0.005); few/no tubule formation (p = 0.03); high nuclear grade (p = 0.038); node positivity (p = 0.037)
Mitotic activity index	***High mitotic activity index:*** Predicts reduced OS* (frequency = 32.5%[138]; both p < 0.05)[137,138]**; reduced RFS* (p = 0.024)[137]**
Mean nuclear area	***High mean nuclear area:*** Predicts reduced OS* (frequency = 50%[128] and 32.5%[138]; both p < 0.05)[128,138]; associated with nuclear atypia (frequency = 32.5%; p = 0.032)[138]; aneuploidy (frequency = 50%; p = 0.01)[128]; high mitotic activity index (frequency = 32.5%; p = 0.011)[138]; high MIB-1 index (frequency = 50%; p = 0.02)[128]; high pathological stage (frequency = 50%; p = 0.01)[128]; high tumour grade (frequency = 50%[128] and 32.5%[138]; both p < 0.05)[128,138]; high PCNA* score (frequency = 50%; p = 0.002)[128]; high AgNOR* quantity (frequency = 50%; p < 0.001)[128]
Standard deviation of nuclear area	***High standard deviation of nuclear area:*** Predicts reduced OS* (frequency = 50%; p = 0.02)[128] and is associated with aneuploidy (frequency = 50%; p = 0.001)[128]; high mitotic activity index (frequency = 32.5%; p = 0.014)[138]; high MIB-1 index (frequency = 50%; p = 0.001)[128]; high tumour grade (frequency = 50%[128] and 32.5%[138]; both p < 0.05)[128,138]; high PCNA* score (frequency = 50%; p < 0.001)[128]; high AgNOR* quantity (frequency = 50%; p < 0.001)[128]; p53 positivity (frequency = 50%; p = 0.005)[128]; Bcl-2 negativity (frequency = 50%; p = 0.04)[128]
Mean nuclear perimeter	***High mean nuclear perimeter:*** Predicts reduced OS* (frequency = 50%; p = 0.01)[128] and is associated with aneuploidy (p = 0.005); high MIB-1 index (p = 0.01); high pathological stage (p = 0.03); high tumour grade (p = 0.002); high PCNA* score (p = 0.001); high AgNOR* quantity (p < 0.001), all at 50% frequency[128]
Standard deviation of nuclear perimeter	***High standard deviation of nuclear perimeter:*** Predicts reduced OS* (frequency = 50%; p = 0.009)[128] and is associated with; aneuploidy (p = 0.001); high MIB-1 index (p = 0.003); high pathological stage (p = 0.001); high

	tumour grade (p = 0.002); high PCNA* score (p = 0.002) ; high AgNOR* quantity (p < 0.001), all at 50% frequency[128]
Nuclear shape factor (Defined as: (4*π*area)/Perimeter2)	***High shape factor:*** Predicts improved OS* (frequency = 42%; both p < 0.05)[128] and is associated with diploidy (p = 0.0007); low MIB-1 index (p = 0.001); low tumour grade (p = 0.0007); p53 negativity (p = 0.005); c-Myc negativity (p = 0.05); low AgNOR* quantity (p = 0.005), all at 42% frequency[128]
Vascular invasion	***High frequency of vascular invasion:*** Associated with ERα/PR negativity (p = 0.0004); high tumour grade (p = 0.035), both at 20% frequency[186]
Tubule formation	***High tubule formation:*** Predicts improved OS* (frequency = 50.5%; p = 0.035)[138]

**MBC: Male Breast Cancer; OS: Overall Survival; RFS: Relapse Free Survival; PCNA: Proliferating Cell Nuclear Antigen; AgNOR: Argyrophillic Nucleolar Organiser Regions; TILs: Tumour Infiltrating Lymphocytes*

***Frequency unavailable from all/some source article(s)*

†Frequency of fibrotic foci >8mm not available from source article

††Surrogate subtype specific breakdown unavailable

Table 5: Novel clusters identified in MBC. Clinical correlations and/or p-values are specified where available.

Category	Cluster	Outcome
Epigenetic	ME1 Epitype (n = 23)[10]	***Associated with:*** Cyclin A positivity (p = 0.012); high fraction of genome alteration (p = 0.0045); high S-phase fraction (p = 0.035); high mitotic activity (p = 1.5e-5) ; luminal M1 transcriptional subgroup[13] ***Compared to the ME2 epitype, ME1 epitype had***: Lower ERα scores (p = 0.048); higher EZH2 expression (p = 3.3e-7); higher activity of proliferation modules (p = 2.8e-7); more frequent hypermethylation of genes involved in epigenetic gene silencing with H3K27me3 (p = 4.4e-153), transcriptional regulation with HOX genes (p = 1.6e-22), cell adhesion pathways (p = 5.6e-5), WNT signalling (p = 2.8e-4), TGF-β signalling (p < 0.001), focal adhesion (p < 0.005), MAPK signalling (p < 0.005), FGFR ligand binding and activation (p < 0.007)
	ME2 Epitype (n = 24)[10]	***Associated with:*** Luminal M2 transcriptional subgroup (p = 0.011)[13]
	Cluster 1 (n = 20)[83]	***Characterised by:*** Hypermethylation of *GSTP1* and *WIF1*; lower methylation levels of *RASSF1A* compared to *MAL*
	Cluster 2 (n = 19)[83]	***Characterised by***: hypermethylation of *GSTP1*
	Cluster 3 (n = 7)[83]	***Characterised by:*** Lower methylation levels of *WIF1* compared to *RASSF1A*; hypermethylation of *RARB* and *GSTP1* and *a*ssociated with *BRCA2* mutation positivity (p = 0.02)
	Cluster 4 (n = 8)[83]	***Characterised by:*** lower methylation levels of *RASSF1A* compared to *TWIST1*
	BRCA2-mutation positive subgroup: Cluster A (n = 12)[83]	***Characterised by:*** Hypermethylation of *GSTP1* and *MAL;* lower *RASSF1A* methylation compared to Cluster B; younger ages of diagnosis compared to other *BRCA2*-mutation positive patients
	BRCA2-mutation positive subgroup: Cluster B (n = 8)[83]	***Characterised by:*** Hypermethylation of *RASSF1A*
Genetic	Luminal M1 (n = 46)[13]	***Associated with:*** HER2 positivity (p = 0.0057); PLAU expression – invasion and metastasis (p = 1.0e-5); AURKA expression – proliferation (p = 0.026)

	Luminal M2 (n = 20)[13]	***Associated with:*** ESR1 expression & ERα positivity (p = 1.3e-8); STAT1 expression – immune response (p = 6.8e-3)
	Male-simple (n = 11)[14]	***Compared to male-complex group, the male-simple group had:*** Lower fraction of altered genome (p = 0.007); lower S-phase fraction (p = 0.02); smaller tumour size (p = 0.004)
	Male-complex (n = 43)[14]	***Characterised by:*** Similarity with the female Luminal B intrinsic subtype; BRCA2 mutation positivity; whole chromosome arm gains
	Cluster A (n = 78)[12]	***Characterised by:*** Partial and whole arm loss of chromosome 16q; higher copy number gain on chromosome 16p compared to Cluster B; higher frequency of loss of chromosome 16q genes compared to Cluster B
	Cluster B (n = 57)[12]	***Characterised by:*** Higher percentage of copy number gain compared to Cluster A; lower frequency of node positivity compared to Cluster A (p = 0.008) and associated with triple negativity (p = 0.042)
	Cluster A (n = 55)[84]	***Characterised by:*** Low rates of copy number gain and amplification.
	Cluster B (n = 51)[84]	***Characterised by:*** Copy number gain in the genes *CCND1, MTDH, CDC6, ADAM9, TRAF4* and *MYC* and independently predicts reduced overall survival (p = 0.009) and associated with high mitotic index (p < 0.001); tumour grade 3 (p = 0.02); large tumour size (p = 0.036)
Transcriptomic	Cluster 1 (n = 41)[11]	**Predicts:** Reduced OS* (p = 0.043) and associated with AURKA signature (proliferation marker) (p = 0.02); HER2 signalling (p = 0.0003); PLAU signature (invasion and metastasis marker) (p = 0.03); STAT1 signature (immune response marker) (p = 0.005)
	Cluster 2 (n = 22)[11]	***Associated with:*** NAT1 upregulation (p = 0.007); CASP3 signature (apoptosis marker) (p = 0.01)
Proteomic	Cluster A1 (Hormone receptor negative) (n = 21)[93]	***Both A1 and A2 clusters:*** Had reduced 5-year overall survival compared to B1 and B2 clusters (p = 0.011) and characterised by ERα negative cases clustering together with PR and AR negative cases; low protein expression of other markers; intermediate histological grade; associated with large tumour size (p = 0.023)
	Cluster A2 (ERα positive high-grade) (n = 37)[93]	***Both A1 and A2 clusters:*** Had reduced 5-year overall survival compared to B1 and B2 clusters (p = 0.011) and characterised by low PR expression; HER2 amplification; high Ki-67 index; accumulation of p21, p16, and p53; expression of EGFR and CK5/6 and associated with: high tumour grade (p = 0.001); high mitotic activity (p < 0.001); node positivity (p = 0.033)
	Cluster B1 (ERα positive intermediate-grade) (n = 34)[93]	***Characterised by:*** Hormone receptor positivity; Bcl-2 and Cyclin D1 positivity; low Ki-67 index; BRST-2 negativity; node negativity

	Cluster B2 (ERα positive low-grade) (n = 37)[93]	***Characterised by:*** Hormone receptor positivity; Bcl-2 and Cyclin D1 positivity; low Ki-67 index; BRST-2 positivity; node positivity
	c-ERβ cluster[123]**	***Characterised by:*** Cytoplasmic expression of both ERβ1 and ERβ2. Also found in FBC*
	PR cluster[123]**	***Characterised by:*** Both PR-A and PR-B isoform action.
	ERα/ERβ/AR cluster[123]**	***Characterised by:*** Collective action of AR with the ER isoforms α, β1, β2, and β5.

**FBC: Female Breast Cancer; OS: Overall Survival*

***breakdown unavailable*

Title: Defining genomic, transcriptomic, proteomic, epigenetic, and phenotypic biomarkers with prognostic capability in male breast cancer: a systematic review

Authors:

[1,2]Subarnarekha Chatterji, MSc (ORCID: 0000-0002-8980-4982)

[1†]Emma Krzoska, MChem (ORCID: 0000-0002-2755-6485)

[1†]Christopher W Thoroughgood, PhD (ORCID: 0000-0002-4470-5411)

[1†]John Saganty, MPhil (ORCID: 0000-0002-3523-227X)

[1,2]Peng Liu, PhD (ORCID: 0000-0002-0058-0434)

[3]Beatrix Elsberger, PhD (ORCID: 0000-0002-2864-5789)

[1,4]Rasha Abu-Eid, PhD (ORCID: 0000-0002-6634-0329)

[1,2*]Valerie Speirs, PhD (ORCID: 0000-0002-0602-4666)

[†]*Authors contributed equally*

Affiliations:

1. School of Medicine, Medical Sciences, and Nutrition, University of Aberdeen, Aberdeen AB25 2ZD, UK
2. Institute of Medical Sciences, University of Aberdeen, Aberdeen AB25 2ZD, UK
3. NHS Grampian, Aberdeen Royal Infirmary, Breast Unit, Foresterhill Rd, Aberdeen AB25 2ZN, UK
4. Institute of Dentistry, University of Aberdeen, Aberdeen AB25 2ZR, UK

*Full Professor and corresponding author: Professor Valerie Speirs

(valerie.speirs@abdn.ac.uk)

Institutional corresponding address:

Institute of Medical Sciences

University of Aberdeen

Foresterhill

Aberdeen AB25 2ZD

United Kingdom

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

All authors agree with the content of the manuscript and consent to publication.

Availability of data and materials

Not applicable. All information can be found in the Appendix~~supplementary files~~ and the reference list.

Competing interests

VS received funding from the University of Aberdeen Development Trust. The other authors declared no conflicts of interest.

Funding

This work was supported by a University of Aberdeen Development Trust Elphinstone Scholarship (SC) and NHS Grampian Endowments (VS).

Author's contributions

Study concept and design: SC, RAE, and VS; Literature search, title screening, and abstract screening: SC; Full-text screening and data extraction: SC, EK, CT, JS, and PL; Accuracy checks: PL, RAE, and VS; Writing – original draft: SC; Writing – review and editing: SC, RAE, BE and VS; Supervision: RAE and VS.

EK, CT, and JS contributed equally. All authors approved the final version to be published.

None of the authors are employed by NIH.

Abstract

While similar phenotypically, there is evidence that male and female breast cancer differ in their molecular landscapes. In this systematic review, we consolidated all existing prognostic biomarker data in male breast cancer, spanning genetics, transcriptomics, proteomics, and epigenetics as well as phenotypic features of prognostic value from articles published in a 29-year period (1992 – 2021). We identified knowledge gaps in the existing literature, discussed limitations of included studies, and outlined potential approaches for translational biomarker discovery and validation in male breast cancer. We also recognised STC2, DDX3, and DACH1 as underexploited markers of male-specific prognostic value in breast cancer. Finally, beyond describing the cumulative knowledge on the extensively researched markers ERα, PR, HER2, AR, and *BRCA2*, we highlighted *ATM*, *CCND1*, *FGFR2*, *GATA3*, HIF1α, *MDM2*, p53 and c-Myc as well-studied predictors of poor survival, that also aligned with several hallmarks of cancer.

Formatted: Font color: Red

Introduction

Breast cancer (BC) affects both sexes but is around 100 times rarer in men[1]. Latest statistics from 2019 show that 25,143 men were affected worldwide, with a 48.1% mortality rate[2]. In comparison, BC affected 1,977,212 women during this period with 34.8% mortality rate[2]. Current clinical management of male breast cancer (MBC) is identical to female breast cancer (FBC), informed by female-only clinical trials. However, MBC differs from FBC in clinical presentation, distribution of histopathological types, and hormone receptor (HR) expression[1,3-5]. Clinical presentation is typically late, MBCs are predominantly oestrogen receptor (ERα) positive (up to 95%), with human epidermal growth factor receptor 2 (HER2) expression uncommon, and triple negativity extremely rare in men[4,6-9].

Formatted: Font color: Red

Hierarchical clustering studies on genetic, transcriptomic, and epigenetic data have identified MBC-specific clusters of prognostic value with limited overlap with the Prediction Analysis of Microarray 50 (PAM50) intrinsic subtypes in FBC[10-15]. Germline mutations in *BRCA2*, established as a high penetrance MBC susceptibility gene have also been extensively researched. Carriers have a lifetime risk of up to 10% of developing cancer, frequently with poor prognosis and aggressive disease characteristics[16-19]. However, despite growing consensus on high-risk men with relevant family history to be offered screening, such an initiative does not yet exist.

Biomarker studies in MBC are few despite rising interest over the past decade. Large scale collaborative studies like the International Male Breast Cancer Program have concentrated mainly on ERα, PR and HER2, which are already integrated into clinical practice[7]. Novel biomarker studies in MBC have revealed numerous candidates with possible male-specific value, but most suffer from small cohorts and lack of independent validation, meaning these remain under-investigated.

While many general reviews on MBC exist, to our knowledge there is no comprehensive systematic review to identify knowledge gaps in MBC biomarkers with prognostic potential. Hence, we exhaustively reviewed molecular studies in MBC adopting a multi-omics and phenotypic approach. We comprehensively describe the existing landscape of prognostic biomarkers in MBC and highlight several molecules that could provide complementary information beyond what is established in BC for future clinical management.

Formatted: Font color: Red

Formatted: Font color: Red

Methods

We conducted and reported this systematic review following Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) recommendations[20].

Search strategy and selection criteria

A systematic search of published literature on MBC biomarkers with a multi-omics and phenotypic approach was conducted using PubMed, Medline, Scopus, Embase, and Web of Science, from the inception of the databases to 16th June 2020. An updated search was performed between 17th June 2020 and 1st November 2021 to include the most recent publications. The representative terms “TITLE (male OR men) AND TITLE (breast OR mammary OR “mammary gland”) AND TITLE (neoplasm OR neoplasia OR malignancy OR malignancies OR cancer OR carcinoma OR tumour OR tumor) AND (KEY (biomarker OR marker)) were used to conduct the electronic search. Complete database specific search terms are detailed in the Appendix (Page 3)~~Supplementary File 1~~.

Inclusion criteria were:

- Primary study population must have included MBC patients and should have been the focus of the study

- Studies must have investigated marker(s) of any omics type or morphological and/or phenotypic features with respect to disease pathogenesis/progression/survival and clinicopathological characteristics of study population(s)

Exclusion criteria were:

- Case reports, case series, letters to the editor, conference abstracts, comments, reviews, and systematic reviews
- Studies conducted on species other than humans
- Original articles in languages other than English
- Primary cohort size ≤ 5

No restrictions were made on methodology, statistical significance of results, or diagnostic/prognostic/predictive value of the biomarkers studied. The selection criteria were intentionally broad to ensure exhaustivity and minimize loss of information. Additionally, reference lists of the included manuscripts were manually searched by SC to identify studies that may have been missed by the electronic search.

Abstracts retrieved from these searches were exported to EndNote referencing software, using which deduplication and screening of titles and abstracts to exclude studies that did not fulfil inclusion criteria was done by SC. Full-text screening of the short-listed articles was conducted in pairs by SC, EK, CT, JS, and PL.

Data extraction

Data extraction of the following variables was performed using Microsoft Excel: first author, published year, country/countries where the study was conducted, study design, method(s), type of tissue tested, cohort size, control group, age (mean/median and range), anatomic stage, histological type and grade, treatment information, St. Gallen classification, nodal status, HR (ERα, PR, HER2) status, number of biomarkers studied, biomarker type (prognostic/predictive/diagnostic), biomarker category (genetic/transcriptomic/proteomic/epigenetic/phenotypic), survival associations, and

associations with clinical features described in each article. ~~with available clinical features~~. FBC data were recorded using the same criteria when present and relevant. To ensure uniformity, all reviewers extracted data from five randomly selected articles for training and calibration. For articles identified in the original search conducted on 16th June 2020, the data extraction process was conducted by two independent reviewers in three pairs (SC + EK, SC + CT, SC + JS). Disagreements were resolved through discussion and with the involvement of a third reviewer when necessary. Data extraction for articles identified in the search from 17th June 2020 to 1st November 2021 was done following the same protocol by SC and PL. Accuracy checks were performed on at least 10% randomly selected articles by RAE and VS.

Quality assessment

Risk of bias assessment was conducted using the Joanna Briggs Institute Critical Appraisal tools using checklists for case-control studies, and analytical cross-sectional studies, as appropriate[21]. Studies had high risk of bias if the response to at least one appraisal question was "No" and/or to multiple questions was "Unclear". If one question had an "Unclear" response, but all other responses were "Yes", the risk of bias was moderate. If the response to all questions was "Yes", the risk of bias was low.

Results

Database search results

In total, 1359 records were retrieved from 5 databases: 306 (PubMed), 576 (Scopus), 187 (Medline), 158 (Embase), 132 (Web of Science). Duplicates (682) were removed, following which 677 articles were screened based on title and abstract. Then, 480 articles were removed as they did not meet the inclusion criteria, leaving 197 articles. These underwent full-text screening, after which 20 articles were removed for not fulfilling the inclusion criteria. Data extraction was performed on the remaining 177 articles. A manual reference search

within included articles revealed 20 relevant articles that were missed by the electronic search. In total, 197 articles were finally included. A PRISMA chart is shown in the Appendix (Page 126)~~Supplementary Figure 1~~.

The included studies were conducted from1992 to 2021. Of these, 27 were descriptive[22-48], and 35 were screening studies[49-82]. Of the latter, 26 reported mutations without any clinical associations. 64 studies[6,7,11,13,15,83-141] reported biomarkers linked to survival and the remaining 78 studies reported biomarkers with clinical associations[10,12-14,49,54,58,59,61,66-68,71,84,87,90,93,106,142-209].

Study Characteristics

We identified 76 case-control studies[10,13-15,22-24,26,31,34,38,44,48,51,54,55,58,60,67,72,73,81,84,88,90,92,94,98,99,106,109,110,120,123,132,136,142-146,148-150,152-160,162,164,166,167,171,172,175-177,182-185,188,191,193,196,202,203,205,206,209,210], of which MBC outcomes were measured against gynaecomastia in 10 studies[23,34,106,132,153,154,158,172,177,193], FBC data in 43 studies[10,13-15,26,31,34,38,48,73,81,84,88,90,92,98,99,109,110,120,123,136,155,159,160,162,164,167,171,175-177,182-185,188,191,196,202,206,209,210], healthy men, women or both in 23 studies[22,24,51,54,55,58,60,67,72,94,142-146,148-150,152,157,175,203,205], and 1st degree male relatives with history of cancer (non-breast) in 1 study[166]. Normal male breast tissue[10,15,44,132,156,162,209], lymph node tissue[156], and non-malignant breast cell lines[10] were used as controls in 7, 1, and 1 study, respectively. Of the case-control studies, 80.3% (n = 61), 5.3% (n = 4), and 14.4% (n = 11) articles had high, moderate, and low risk of bias, respectively (Appendix Page 5~~Supplementary Table 1~~).

The remaining 121 studies were cross-sectional[6,7,11,12,25,27-30,32,33,35-37,39-43,45-47,49,50,52,53,56,57,59,61-66,68-71,74-80,82,83,85-87,89,91,93,95-97,100-105,107,108,111-119,121,122,124-128,130,131,133-135,137-141,147,161,163,165,168-170,173,174,178-181,186,187,189,190,192,194,195,197-201,204,207,208,211,212]. Most had MBC patients as their sole cohort, while 2 studies included FBC patients with MBC-affected relatives alongside their primary MBC cohort[76,79]. Out of the cross-sectional studies, 56.2% (n = 68) and 43.8% (n =

53) had high and low risks of bias, respectively (Appendix Page 11~~Supplementary Table 1~~). Study characteristics are summarized in the Appendix (Page 19)~~Supplementary Table 2~~.

We identified 304 biomarkers in total and classified them according to their respective omics/phenotypic categories. The 10 most studied biomarkers from each category, based on the number of reporting studies and associations with clinical features are detailed in Tables 1-4. The full list of biomarkers with their clinical associations, and all reported pathological gene variations are provided in the Appendix (Page 43-125)~~Supplementary Tables 3-7~~.

Proteomic markers

ERα, PR, and HER2

These receptors currently define standard-of-care in BC and were studied both as biomarkers and clinical factors associated with other biomarkers. The MBC cohorts studied were overwhelmingly ERα-positive, predicting improved OS and DFS[7,123], while ERα-negativity, predicted reduced OS[104,118,122,134] and younger age of diagnosis[93]. Like FBCs, PR was frequently co-expressed with ERα, its positivity mostly predicting prognostic benefit[7,87,93,104,105,118,122].

Overexpression and amplification of HER2 was evaluated by immunohistochemistry (IHC) and fluorescent *in-situ* hybridisation (FISH), the latter being detailed in the genetics/transcriptomics markers section. Overexpression was associated with aggressive features and reduced survival by every study investigating HER2 prognostic value[6,87,95,101,129,188,198] (Table 1).

St Gallen surrogate classification

Luminal B and triple negative MBCs had poor survival and aggressive features[87,101,119,190,208], with the latter more frequent in men of black ethnicity[101]. Basal-like MBCs were diagnosed at younger age than Luminal A/B MBCs[190]. Several biomarkers were expressed differentially between the Luminal classifications. GCDFP15-positivity[187] and p53-negativity[181] were

associated with Luminal A MBCs, while ATF3, FATP1, p21-positivity, and Bcl2-negativity were associated with HER2-negative Luminal B MBCs[93,100]. The latterformer also had higher expression of EGFR and NF-κB compared to Luminal A MBCs[37] (Appendix Page 43Supplementary Table 5).

Other proteomic markers

AR expression had both prognostic advantage[6,7,116,123,131,179,200] and disadvantage[94,96,117]. Interestingly, two out of three studies predicting poor outcome were conducted on ethnically homogeneous Chinese populations[94,117]. Like FBCs, AR was consistently co-expressed with ERα[94,116,131,133,179]. AR co-expression with ERα and FOXA1 predicted improved OS[123] and DFS[6], respectively.

High tumour proliferation index (represented by Ki-67/MIB1 index) consistently predicted poor survival and aggressive disease[87,93,113,115,118,129,131,133,135,184,186,196,197].

Of the most studied markers, p53[93,119,128,129,131], p21[93,125,160,196], EGFR[118,188,190] and c-Myc[125,129] predicted reduced survival. The tumour hypoxia markers HIF1-α, CA-9 and Glut-1 along with their co-expression profiles also predicted poor outcome[124,141,180].

Relatively few biomarkers predicted improved outcome and were rarely reported by multiple studies. Bcl-2[93,181,189,194,202] and Cyclin D1 positivity[93,121,125,133] were mostly linked to improved outcome.

Several markers displayed sex-specific differences in expression. Hormone receptors ERα[185], PR[202], AR[123], ERβ1[123] and ERβ2[123] were expressed more frequently in MBCs than FBCs. STC2[109], IGF1-R[188], CAXII[188], p21[160,196], p27[196], p53[160] and Bcl-2[202] were also overexpressed in MBC compared to FBC, while the opposite was true for DACH1[182], PD-1[183], MET[188], FGFR2[188], CD44v6[188] and GATA3[120]. DDX3 had higher cytoplasmic expression but lower nuclear expression in MBCs compared to FBCs[102]. Improved survival or favourable outcomes in MBC were linked to STC2[109], p27[125,196,197], Bcl-2[93,181,189], and high cytoplasmic DDX3 expression[102]. The opposite was true for p21[93,125],

Formatted: Font color: Red

p53[31,93,119,128,129,131,160,181,202], DACH1[182], and GATA3[90,120]. The prognostic value of STC2[109], DDX3[102], and DACH1[182] were assessed by only one study each (Table 1 and Appendix Page 43~~Supplementary Table 5~~).

Genetic and transcriptomic markers

Pathogenic variations in BRCA genes with prognostic value

Germline *BRCA2* mutations are the most frequently reported pathological gene variations in MBC. These predicted reduced overall (OS), disease-free (DFS), and disease-specific survival (DSS)[85,87,96], and aggressive features like young age of diagnosis, bilaterality, contralaterality, node positivity, advanced tumour grade, ERα/PR-negativity, HER2-positivity, high Ki-67 index, personal history of cancer[59,61,68,87,149,164,167,170,173,175], high frequency of genetic aberrations[175], amplifications[88] and copy number variations (CNV)[168] of several cancer-related genes. *BRCA2* mutations were more frequent and had more aggressive features in MBCs compared to FBCs[59,164]. In contrast, germline *BRCA1* mutations were less frequent in MBCs[59] and had less pronounced prognostic value, with links to advanced tumour grade[164], ERα-negativity[170], and family history of pancreatic cancer[66] (Table 2).

Germline mutations were most frequently reported in *BRCA2* and *BRCA1* (28 and 12 studies, respectively), followed by *CHEK2, PALB2*, and *ATM* (9, 7, and 3 studies respectively).

Pathogenic variations in other genes with prognostic value

While uncommon in MBC (0 - 9% of all cases[6,7,123]), HER2 amplification predicted reduced OS, younger age of diagnosis, large tumour size, advanced disease stage, and both regional and distant metastasis[84,86,93,95].

Several genetic variations predicted reduced OS. These included somatic mutations in *PIK3CA*[88], *GATA3*[90] and *THY1*[92], and amplifications in *MDM2, PAK1, TGFB2, SCYL3*[88], *CCND1* and *EMSY*[84]. Mutations in DNA repair genes were enriched in Luminal A-like MBCs

Formatted: Font color: Red

Formatted: Font color: Red

compared to matched FBCs and predicted reduced survival in general[90]. In contrast, survival benefit was associated with relatively few genetic/transcriptomic variations, with only upregulation of miR-125b, which targets genes covering multiple biological signalling pathways in many cancers[213], being reported in >1 study[177,209] (Table 2 and Appendix Page 71~~Supplementary Table 3~~).

Pathogenic variations associated with MBC risk

Germline mutations in *PALB2* and *RAD51D*[54] had the highest odds-ratios (17.30, 8.58; 11.20, 10.18 , using the Exome Variant Server and Non-Finnish European datasets, respectively), followed by *MUTYH* (4.54)[147], *CHEK2* (4.47)[58], and *SULT1A1* (3.09; A/A polymorphism)[148]. Copy number (CN) gain in *PALB2* was associated with node negativity[12] and its mutated status was associated with bilaterality[49]. Increased MBC risk was also linked to single nucleotide polymorphisms (SNPs) in multiple genes, with rs3803662 (*TOX3*) reported by two independent groups[144,145].

Screening studies from 1995 to 2021 identified pathogenic mutations in several genes in MBC, most of them germline. The *CHEK2* c.1100delC mutation was reported most frequently[49,52,54,58,63,66], followed by the *BRCA2* c.6174delT[57,61,64,66] and c.771_775delTCAAA (also known as c.999del5)[59,69,72,81] (Appendix Page 100~~Supplementary Table 4~~).

Epigenetic markers

Advanced tumour grade, high mitotic index, large tumour size, ERα-negativity, and mutated *BRCA2* were linked to promoter hypermethylation of most reported genes[83,155,156]. Interestingly, hypermethylated *RASSF1A* and *RARB* were linked to both ERα-negativity and PR-positivity, which have opposing clinical significance in FBC[157]. Hypermethylated *RASSF1A* was also linked to HER2-positivity[156]. High methylation indices, high methylation rate, and high number of methylated genes predicted reduced OS and DSS, and aggressive features like *BRCA2*-mutation, high mitotic index, high tumour grade, and large tumour

Formatted: Font color: Red

size[15,83]. Only one study associated promoter hypermethylation of any gene to survival, with hypermethylated *TWIST1* predicting reduced DSS, especially in *BRCA2*-mutated MBCs[83].

Conflicting results were reported on *AR* promoter hypermethylation. Virtually non-existent *AR* methylation and very little methylation of its co-regulators was observed in MBC when compared to gynaecomastia [154]. However, tumour DNA had higher *AR* methylation compared to normal tissue and lymph nodes (both patient unmatched)[156]. *AR* hypermethylation was also associated with wild type *BRCA1/2*[156].

Regarding sex-specific epigenetic differences, reduced methylation levels were more common in both invasive carcinoma (IC) and ductal carcinoma *in-situ* adjacent to invasive carcinoma (DCIS-AIC) in MBC compared to FBC. Only *GATA5, THBS1, MSH6*, and *RASSF1A* were more heavily methylated in males compared to females[155,157].

Within MBC cohorts, higher methylation was reported in DCIS-AIC compared to pure ductal carcinoma *in-situ* (DCIS), while IC had higher methylation levels compared to DCIS-AIC. Hypermethylation in normal breast tissue and lymph nodes (both patient unmatched) was consistently less frequent compared to IC[156] (Table 3 and Appendix Page 113~~Supplementary Table 6~~).

Morphological and/or phenotypic features

Several morphological features of MBC had prognostic significance. Unsurprisingly, high mitotic activity index predicted reduced survival[137]. High nuclear area and high variation in nuclear size predicted poor survival and aggressive features [128,138]. Presence of fibrotic foci predicted reduced OS[124,137] and recurrence-free survival (RFS)[137], and advanced tumour grade, nodal involvement, and low tubule formation[124]. The latter also predicted reduced OS[138]. Like FBCs, low density of tumour infiltrating lymphocytes (TILs) predicted reduced OS and RFS[137], and nodal involvement[186]. Intriguingly, HER2-positive MBCs had higher density of TILs than HER2-negative MBCs, although HER2 overexpression predicted poor prognosis[137].

Low grade ERα-positive MBCs had reduced elastosis than matched FBCs. In FBCs elastosis is strongly associated with ERα expression. Therefore, low frequency of elastosis in MBC despite overwhelming ERα-positivity suggests sex-specific ERα action[206]. Morphological features of both lymphangiogenesis and angiogenesis like high lymphatic vessel density, high distribution of lymphatic vessels, and high frequency of vascular invasion were linked to advanced tumour grade, high tumour proliferation index, and hormone receptor negativity, albeit without reproduction[186]. In agreement, high CD34 expression representing microvascular density predicted reduced RFS and advanced disease stage[130] (Table 4 and Appendix Page 119~~Supplementary Table 7~~).

Novel subgroups in MBC

The first major hierarchical clustering study identifying male-specific BC subgroups was done by Johansson et al[13]. Luminal M1 group exhibited HER2-positivity and associated with invasion, proliferation, and metastasis, while Luminal M2 group displayed ERα-positivity and associated with anti-tumour immune response[13]. They also previously identified Male-simple and Male-complex clusters. The former was genetically stable and differed from female intrinsic subtypes, while the latter consisted of *BRCA2*-mutated MBCs, with worse prognosis and genetic overlap with the Luminal B intrinsic type[14].

These results were validated by a genome-wide methylation study revealing two stable MBC epitypes (ME1 and ME2)[10]. ME1 epitype displayed high mitotic activity, high fraction of genome alteration, Cyclin A-positivity, and ERα-negativity, and frequent hypermethylation of genes involved in key pathways (H3K27me3 epigenetic silencing, transcriptional regulation with HOX genes, WNT, TGF-β, and MAPK signalling, cellular and focal adhesion, and FGFR ligand binding and activation). ME1 and ME2 epitypes aligned with the Luminal M1 and M2 subgroups, respectively[13].

A later study reported 4 epigenetics-based clusters based on the relative promoter hypermethylation levels of *RASSF1A, GSTP1, WIF1, RARB,* and *MAL*. Notably, Cluster 3

associated with mutated *BRCA2* (p = 0.02)[83]. This study performed a subgroup analysis on *BRCA2*-mutated MBCs which separated into 2 clusters based on the hypermethylation levels of *GSTP1, MAL,* and *RASSF1A*[83].

Most recently, two clusters were reported based on RNASeq data[11]. Cluster 1 had reduced OS and associated with HER2 signalling, proliferation, invasion and metastasis, and immune response, while Cluster 2 associated with the apoptosis hallmark and NAT1 signalling[11]. These clusters had limited overlap with the Luminal M1 and M2 subgroups. Immune response clustered with invasion and metastasis, and proliferation, directly contradicting Luminal M1 and M2 characteristics[11,13].

Cluster separation was also reported based on chromosome 16q CNVs. Cluster A had low rates of CN gain and amplification, predicting prognostic benefit, while Cluster B had aggressive features[84]. Building on this work, another study reported clusters based on chromosome 16q CNVs, where Cluster A associated with node positivity, and Cluster B with triple negativity[12].

Four clusters based on immunohistochemical markers were described[93]. Clusters A1 and A2 had aggressive characteristics; A1 defined by hormone negativity, and A2 by ERα-positivity, PR-negativity, and HER2-amplification. The less aggressive clusters B1 and B2 were histologically identical, although B1 exhibited BRST-2 positivity and nodal involvement, while B2 had the opposite features[93].

MBC clusters separating on ER/PR isoforms were also reported[123]. These respectively separated on the cytoplasmic expression of ERβ1 and 2, PR isoforms A and B, and collective action of AR with ERα and β1 isoforms. Only cytoplasmic-ERβ cluster had FBC overlap[123] (Table 5).

Alignment of biomarkers with the Hallmarks of Cancer

Formatted: Font: Bold

Upon interrogation of the COSMIC database[214], certain genetic, transcriptomic, proteomic, or epigenetic markers aligned with the 2000 and 2011 Hallmarks of Cancer[215,216]. These had

prognostic impact in MBC and/or differential expression between the sexes. Certain molecules identified in the same categories were also speculatively linked to the most recent Hallmarks of Cancer[217] (both described on page 127 of the Appendix~~in Supplementary Figure 2~~). Based on these associations, these molecules may warrant further research: *ATM*, *CCND1* (Cyclin D1), *GATA3*, *FGFR2*, *HIF1A* (HIF1-α), *MDM2*, *MYC* (c-Myc), and *TP53* (p53). These were linked to multiple hallmarks of cancer through promoter and/or suppressor action, were associated with ≥1 clinical feature across multiple omics categories and could predict survival in at least one of these categories.

Discussion

MBC is receiving increased recognition. A bibliometric analysis revealed that most publications in MBC focused on clinical risk factors and management, followed by comparisons against FBC[218]. MBC management is still largely defined by superficial extrapolation of FBC standard-of-care despite mounting evidence of sex-related differences. Recognising a need to identify translationally valuable biomarkers that can define a male-inclusive picture of BC, this systematic review comprehensively described the biomarker landscape of MBC and identified markers that may aid future clinical management. To our knowledge, this is the first exhaustive systematic review on the subject.

ERα and PR emerged as having sex-specific regulatory characteristics. Although a known modulator of ERα binding in FBC, many PR binding sites were devoid of ERα in MBC[98]. Hierarchical clustering studies found independent PR clusters[123] in MBC, while ERα/PR action clustered together in FBC[98,123]. Mathematical modelling revealed no continuous dependency effect on ERα for PR[31]. Furthermore, two FBC clusters were identified based on PR action in FBC but not in MBC[171].

Regarding ER isoforms, ERα/ERβ/AR[123], and ERα/FOXA1/AR coaction predicted improved survival in MBC[6]. As most ERα binding sites in both sexes are independent of FOXA1[98], this

suggests an intermediary role of FOXA1 (and possibly ERβ) in ERα/AR interaction in MBC. This requires elucidation.

AR expression, when studied independently, predicted contradicting prognostic outcomes[6,7,94,96,116,117,123,131,179,200]. Epigenetic findings on AR were also inconsistent. AR hyperactivity in ERα-positive MBC was speculated based on hypomethylation of AR and its co-regulators compared to gynaecomastia[154], while another study demonstrated AR hypermethylation in tumours compared to unmatched normal lymph nodes and breast tissue[156]. Therefore, the exact impact of AR methylation remains unclear. The contradictory role of AR was further highlighted by its value as a therapeutic target in MBC. Phase II trial data showed that the AR inhibitor enzalutamide was well-tolerated in both sexes, and improved PFS in both HR positive and androgen-driven triple negative BC[219,220]. Similar results were seen with the AR/CYP17-L inhibitor seviteronel in both sexes[221]. In FBC, AR plays a compensatory role for ERα in ERα-negative/AR-positive FBC, and this is supported by overlapping binding characteristics of ERα and AR[98,222]. However, the same cannot be speculated for MBC as most patients are ERα/AR-positive. A partial explanation is offered by the sex-specific nature of prognostic ability of ERα binding sites[98], but we await a complete picture of ERα/AR interaction in MBC. Intriguingly, AR-driven tumour-suppressor activity was observed in ERα/AR-positive BC cell lines and FBC patient-derived explant (PDE) models, clearly supporting agonism over antagonism of AR as a more valuable treatment strategy[223].

The aggressive nature of germline *BRCA2* mutations has been established in MBC [59,61,68,87,149,164,167,170,173,175]. However, *BRCA2* is yet to inform clinical management, despite there being an argument for male patients with family history of *BRCA2*-related cancers (breast, ovarian, prostate, and pancreatic) to be screened and offered genetic counselling[224]. The incidence of *BRCA2*-mutated MBCs in different ethnicities also need to be established.

Given the negative prognostic effect of somatic mutations in the *PIK3CA* gene in MBC[88,158], the SOLAR-1 trial is worth mentioning. This randomised phase-3 trial included men and

postmenopausal women with HR-positive/HER2-negative BC with mutated *PIK3CA* and demonstrated improved OS when the PI3KA-specific inhibitor Alpelisib was administered with Fulvestrant[225]. This trial is an encouraging example of positive advances being made towards inclusion of men in clinical trials.

Discovery of novel markers in MBC has historically suffered due to small cohort sizes and lack of prospective validation. This generally aligns with the broader picture of biomarker discovery in oncology, where most molecules are often left unexplored beyond their initial identification and establishment of a significant survival association. The relative rarity of MBC and small number of research papers brings this into sharp focus.

As shown in the Appendix (Page 129)~~Supplementary Figure 3~~, most of the well-studied biomarkers with hallmarks functions also regulate the G1/S phase transition pathway of the cell cycle along with *RB1, MDM2, ATR, CHEK2, CDKN1A* (p21)*, CDKN1B* (p27)*, CDKN2A,* and *CCNE1*, alterations of which were also linked with MBC clinical outcome in at least one -omics category (~~Supplementary~~ Figure 14). Most of these biomarkers predicted poor survival, which justifies focused drug-target identification studies through selective inhibition of regulatory pathways. The role of Cyclin D1 is especially worth investigating, as it predicted improved survival as a proteomic marker[93,121,125,133], but the opposite as a genetic marker (*CCND1*)[84].

In this regard, the CDK4/6 inhibitor Palbociclib was approved for use in metastatic MBC[226]. Literature supporting the use of CDK4/6 inhibitors in combination with tamoxifen/AI and GnRH in a metastatic setting also exist[227,228]. A recent case report described complete remission of a metastatic MBC patient following treatment with Abemaciclib, Fulvestrant, and Leuprolide[229]. The evidence gathered here supports this approach. However, extending this to the adjuvant setting for MBC may be premature based on results of the PALLAS trial[230].

Amongst the plethora of molecules we identified, STC2[109], DDX3[102], and DACH1[182] are especially worth highlighting in those that were only reported in single studies. STC2 is

involved in pathways regulating stress response, hypoxia, apoptosis prevention, cellular proliferation, migration, and immune response[231]. Tumour and stromal STC2 expression were observed in some 50% and 65% of MBC patients, respectively[109]. DDX3 promotes cancer progression by remodelling the tumour microenvironment[232]. Nuclear and cytoplasmic expression of DDX3 was observed in 42.5% and 20.8% of MBC patients, respectively[102]. DACH1 is a tumour suppressor implicated in the inhibition of invasion and metastasis via downregulation of matrix metalloproteinase 9 transcription, whose positivity was observed in 35.7% MBC cases[182,233]. These proteins were differentially expressed between the sexes and could predict survival in MBC, however, remains underexploited from a translational perspective.

Defining morphological markers of prognosis is necessary as these can be the primary diagnostic considerations. Variation in nuclear area and size are obvious markers of negative prognosis in MBC, which was confirmed in two studies we reviewed[128,138]. The presence/dimensions of fibrotic foci emerged as important markers predicting reduced survival[124,137]. Suggested to be the link between hypoxia and aggressive tumour characteristics, these results were validated by the unfavourable prognostic value of the hypoxia markers HIF1-α, CA-9, and Glut-1[124,141].

Ethnic homogeneity may explain lack of reproducibility for certain studies, such as conflicting prognostic impact for certain markers. This is concerning, as US data show that the age-standardized incidence of MBC in non-Hispanic black men is 2.6 times higher than their white counterparts for ERα-positive/HER2-negative BC[234]. Despite this, no molecular studies investigating ethnicity-specific differences in MBC exist, leaving a significant knowledge gap. Also, ethnicities were not specified in the clustering studies, and therefore no conclusions could be drawn regarding their global representation.

The appropriate selection of controls is another area that may require future consideration. For example, some studies used gynaecomastia samples as controls, as normal male breast tissue is difficult to obtain. However, gynaecomastia is now treated as being aetiologically

distinct from MBC and therefore unlikely to be a suitable comparison[235,236] presenting potential limitations.

Conclusion

Our results demonstrate MBC is a heterogeneous and complex condition with striking distinctions from FBC. MBC research has seen remarkable evolution, from simply replicating FBC marker studies, to its treatment as a separate condition with exploratory studies contributing to a male-specific molecular profile.

We identified conflicting evidence regarding regulation, expression, and prognostic utility of key BC markers alongside sex-specific differences. Considering this, the role of ERα, PR, and AR need to be re-established in a male-specific setting. Developing suitable MBC laboratory models are necessary to achieve this. Beyond the established BC markers, we highlighted that STC2, DDX3, and DACH1 may have grounds for further investigation. We also identified *ATM*, *CCND1* (Cyclin D1), *FGFR2*, *GATA3*, *HIF1A* (HIF1-α), *MDM2*, *MYC* (c-Myc) as well studied predictors of poor prognosis.

To effectively drive the inclusion of male-specific biomarkers from bench to clinical practice, inclusion of men in randomized clinical trials is crucial. Positive advances have been made in this respect with the International Male Breast Cancer Program making a concerted effort to run male-specific trials, and at least two MBC phase-II trials investigating GnRH/AI/tamoxifen and AR-antagonists being reported[221,237,238] alongside the SOLAR-1 trial discussed above[225].

Comprehensively defining biomarkers of translational value adopting a multi-omics and phenotypic approach alongside complementary image analysis studies harnessing modern spatial biology techniques that combine artificial intelligence and digital pathology could yield high-quality spatially resolved molecular profiles of MBC, improving our understanding of this rare cancer.

References:

We cited 239 references in this manuscript, including the 197 studies that met the inclusion criteria of the systematic review. The first 100 references are listed below with the rest in the Appendix (Page 130).

1. Fox S, Speirs V, Shaaban AM. Male breast cancer: an update. *Virchows Arch* 2022; **480**(1): 85-93.
2. Xu S, Liu Y, Zhang T, et al. The Global, Regional, and National Burden and Trends of Breast Cancer From 1990 to 2019: Results From the Global Burden of Disease Study 2019. *Front Oncol* 2021; **11**: 689562.
3. Ferzoco RM, Ruddy KJ. The Epidemiology of Male Breast Cancer. *Curr Oncol Rep* 2016; **1**(18): 1-6.
4. Gucalp A, Traina TA, Eisner JR, et al. Male breast cancer: a disease distinct from female breast cancer. *Breast Cancer Res Treat* 2019; **173**(1): 37-48.
5. Darkeh MHSE, Azavedo E. Male breast cancer clinical features, risk factors, and current diagnostic and therapeutic approaches. *Int J Clin Med* 2014; **5**(17): 1068-86.
6. Humphries MP, Rajan SS, Honarpisheh H, et al. Characterisation of male breast cancer: a descriptive biomarker study from a large patient series. *Sci Rep* 2017; **7**(1): 1-9.
7. Cardoso F, Bartlett JMS, Slaets L, et al. Characterization of male breast cancer: Results of the EORTC 10085/TBCRC/BIG/NABCG International Male Breast Cancer Program. *Ann Oncol* 2018; **29**(2): 405-17.
8. Cruz-Tapias P, Rubiano W, Rondón-Lagos M, Villegas VE, Rangel N. Intrinsic Subtypes and Androgen Receptor Gene Expression in Primary Breast Cancer. A Meta-Analysis. *Biology* 2021; **10**(9): 834.
9. Rozeboom B, Dey N, De P. ER+ metastatic breast cancer: past, present, and a prescription for an apoptosis-targeted future. *Am J Cancer Res* 2019; **9**(12): 2821-31.
10. Johansson I, Lauss M, Holm K, et al. Genome methylation patterns in male breast cancer - Identification of an epitype with hypermethylation of polycomb target genes. *Mol Oncol* 2015; **9**(8): 1565-79.
11. Zelli V, Silvestri V, Valentini V, et al. Transcriptome of Male Breast Cancer Matched with Germline Profiling Reveals Novel Molecular Subtypes with Possible Clinical Relevance. *Cancers* 2021; **13**(18): 4515.
12. Lacle MM, Kornegoor R, Moelans CB, et al. Analysis of copy number changes on chromosome 16q in male breast cancer by multiplex ligation-dependent probe amplification. *Mod Pathol* 2013; **26**(11): 1461-7.

13. Johansson I, Nilsson C, Berglund P, et al. Gene expression profiling of primary male breast cancers reveals two unique subgroups and identifies N-acetyltransferase-1 (NAT1) as a novel prognostic biomarker. *Breast Cancer Res* 2012; **14**(1): 1-5.
14. Johansson I, Nilsson C, Berglund P, et al. High-resolution genomic profiling of male breast cancer reveals differences hidden behind the similarities with female breast cancer. *Breast Cancer Res Treat* 2011; **129**(3): 747-60.
15. Kornegoor R, Moelans CB, Verschuur-Maes AHJ, et al. Promoter hypermethylation in male breast cancer: Analysis by multiplex ligation-dependent probe amplification. *Breast Cancer Res* 2012; **14**(4): 1-9.
16. Barnes DR, Silvestri V, Leslie G, et al. Breast and Prostate Cancer Risks for Male BRCA1 and BRCA2 Pathogenic Variant Carriers Using Polygenic Risk Scores. *J Natl Cancer Inst* 2022; **114**(1): 109-122.
17. Fentiman IS. Male breast cancer is not congruent with the female disease. *Crit Rev Oncol Hematol* 2016; **100**(101): 119-24.
18. Rizzolo P, Silvestri V, Tommasi S, et al. Male breast cancer: genetics, epigenetics, and ethical aspects. *Ann Oncol* 2013; **24**(Supplement 8): viii75-viii82
19. Fentiman I. Male breast cancer: a review. *Ecancermedicalscience* 2009; **3**: 140.
20. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *Syst Rev* 2021; **10**(1): 1-11.
21. Moola S, Munn Z, Tufanaru C, et al. Chapter 7: Systematic reviews of etiology and risk. In: Aromataris E, Munn Z, eds. *JBI Manual for Evidence Synthesis*. JBI, 2020.
22. Rizzolo P, Silvestri V, Valentini V, et al. Evaluation of CYP17A1 and CYP1B1 polymorphisms in male breast cancer risk. *Endocr Connect* 2019; **8**(8): 1224-9.
23. Di Oto E, Biserni GB, Varga Z, et al. X chromosome gain is related to increased androgen receptor expression in male breast cancer. *Virchows Arch* 2018; **473**(2): 155-63.
24. Karhu R, Laurila E, Kallioniemi A, Syrjäkoski K. Large genomic BRCA2 rearrangements and male breast cancer. *Cancer Detect Prev* 2006; **30**(6): 530-4.
25. Rudlowski C, Friedrichs N, Faridi A, et al. HER-2/neu gene amplification and protein expression in primary male breast cancer. *Breast Cancer Res Treat* 2004; **84**(3): 215-23.
26. Wick MR, Sayadi H, Ritter JH, Hill DA, Reddy VB, Gattuso P. Low-stage carcinoma of the male breast. A histologic, immunohistochemical, and flow cytometric comparison with localized female breast carcinoma. *Am J Clin Pathol* 1999; **111**(1): 59-69.
27. Prechtel D, Werenskiold AK, Prechtel K, Keller G, Hofler H. Frequent loss of heterozygosity at chromosome 13q12-13 with BRCA2 markers in sporadic male breast cancer. *Diagn Mol Pathol* 1998; **7**(1): 57-62.

28. Wingren S, van den Heuvel A, Gentile M, Olsen K, Hatschek T, Soderkvist P. Frequent allelic losses on chromosome 13q in human male breast carcinomas. *Eur J Cancer* 1997; **33**(14): 2393-6.
29. Chuaqui RF, Sanz-Ortega J, Vocke C, et al. Loss of Heterozygosity on the Short Arm of Chromosome 8 in Male Breast Carcinomas. *Cancer Res* 1995; **55**(21): 4995-8.
30. Biserni GB, Di Oto E, Moskovszky LE, Foschini MP, Varga Z. Preferential expression of NY-BR-1 and GATA-3 in male breast cancer. *J Cancer Res Clin Oncol* 2018; **144**(2): 199-204.
31. Kornegoor R, van Diest PJ, Buerger H, Korsching E. Tracing differences between male and female breast cancer: both diseases own a different biology. *Histopathology* 2015; **67**(6): 888-97.
32. Verschuur-Maes AH, Kornegoor R, de Bruin PC, Oudejans JJ, van Diest PJ. Do columnar cell lesions exist in the male breast? *Histopathology* 2014; **64**(6): 818-25.
33. Sanchez-Munoz A, Roman-Jobacho A, Perez-Villa L, et al. Male Breast Cancer: Immunohistochemical Subtypes and Clinical Outcome Characterization. *Oncology* 2012; **83**(4): 228-33.
34. Kraus TS, Cohen C, Siddiqui MT. Prostate-specific antigen and hormone receptor expression in male and female breast carcinoma. *Diagn Pathol* 2010; **5**(1): 1-6.
35. Shah P, Robbani I, Shah O. Clinicopathological study of male breast carcinoma: 24 Years of experience. *Ann Saudi Med* 2009; **29**(4): 288-93.
36. Jamal S, Mushtaq H, Mubarik A, Malik TM. Estrogen Receptor, Progesterone Receptor, HER2/neu, P53 and Ki-67 status of Male Breast Carcinomas in Pakistan. *Asian Pac J Cancer Prev* 2009; **10**(6): 1067-70.
37. Ge Y, Sneige N, Eltorky MA, et al. Immunohistochemical characterization of subtypes of male breast carcinoma. *Breast Cancer Res* 2009; **11**(3): 1-8.
38. Milias S, Kalekou H, Bobos M, et al. Immunohistochemical investigation of CD34 antigen in male breast carcinoma. *Clin Exp Med* 2007; **7**(3): 122-6.
39. Ciocca V, Bombonati A, Gatalica Z, et al. Cytokeratin profiles of male breast cancers. *Histopathology* 2006; **49**(4): 365-70.
40. Frangou EM, Lawson J, Kanthan R. Angiogenesis in male breast cancer. *World J Surg Oncol* 2005; **3**(1): 1-4.
41. Murphy CE, Carder PJ, Lansdown MRJ, Speirs V. Steroid hormone receptor expression in male breast cancer. *Eur J Surg Oncol* 2006; **32**(1): 44-7.
42. Kidwai N, Gong Y, Sun X, et al. Expression of androgen receptor and prostate-specific antigen in male breast carcinoma. *Breast Cancer Res* 2003; **6**(1): 1-6.
43. Blin N, Kardaś I, Welter C, et al. Expression of the C- erbB2 proto-oncogenein male breast carcinoma: Lack of prognostic significance. *Oncology* 1993; **50**(6): 408-11.

44. Kardaś I, Seitz G, Limon J, et al. Retrospective analysis of prognostic significance of the estrogen-inducible pS2 gene in male breast carcinoma. *Cancer* 1993; **72**(5): 1652-6.
45. Alm P, Alumets J, Bak-Jensen E, Olsson H. Neuroendocrine differentiation in male breast carcinomas. *Apmis* 1992; **100**(7-12): 720-6.
46. Li X, Wang J, Fu J, Xie X, Xie X. Prognostic significance of preoperative serum inflammation markers in patients with male breast cancer. *Transl Cancer Res* 2021; **10**(9): 4002-8.
47. Hartkopf AD, Taran FA, Walter CB, et al. Detection and prevalence of disseminated tumor cells from the bone marrow of early stage male breast cancer patients. *Breast Cancer Res Treat* 2015; **152**(1): 51-5.
48. Bloom KJ, Govil H, Gattuso P, Reddy V, Francescatti D. Status of HER-2 in male and female breast carcinoma. *Am J Surg* 2001; **182**(4): 389-92.
49. Szwiec M, Tomiczek-Szwiec J, Kluźniak W, et al. Genetic predisposition to male breast cancer in Poland. *BMC Cancer* 2021; **21**(1): 1-8.
50. Ben Kridis-Rejeb W, Ben Ayed-Guerfali D, Ammous-Boukhris N, et al. Identification of novel candidate genes by exome sequencing in Tunisian familial male breast cancer patients. *Mol Biol Rep* 2020; **47**(9): 6507-16.
51. Vietri MT, Caliendo G, D'Elia G, et al. BRCA and PALB2 mutations in a cohort of male breast cancer with one bilateral case. *Eur J Med Gen* 2020; **63**(6): 103883.
52. Tedaldi G, Tebaldi M, Zampiga V, et al. Male Breast Cancer: Results of the Application of Multigene Panel Testing to an Italian Cohort of Patients. *Diagnostics* 2020; **10**(5): 269.
53. Scarpitta R, Zanna I, Aretini P, et al. Germline investigation in male breast cancer of DNA repair genes by next-generation sequencing. *Breast Cancer Res Treat* 2019; **178**(3): 557-64.
54. Rizzolo P, Zelli V, Silvestri V, et al. Insight into genetic susceptibility to male breast cancer by multigene panel testing: Results from a multicenter study in Italy. *Int J Cancer* 2019; **145**(2): 390-400.
55. Silvestri V, Rizzolo P, Zelli V, et al. A possible role of FANCM mutations in male breast cancer susceptibility: Results from a multicenter study in Italy. *Breast* 2018; **38**: 92-7.
56. Fostira F, Saloustros E, Apostolou P, et al. Germline deleterious mutations in genes other than BRCA2 are infrequent in male breast cancer. *Breast Cancer Res Treat* 2018; **169**(1): 105-13.
57. Schayek H, Korach H, Laitman Y, Bernstein-Molho R, Friedman E. Mutational analysis of candidate genes in Israeli male breast cancer cases. *Breast Cancer Res Treat* 2018; **170**(2): 399-404.

58. Hallamies S, Pelttari LM, Poikonen-Saksela P, et al. CHEK2 c.1100delC mutation is associated with an increased risk for male breast cancer in Finnish patient population. *BMC Cancer* 2017; **17**(1): 1-5.
59. Deb S, Jene N, investigators K, Fox SB. Genotypic and phenotypic analysis of familial male breast cancer shows under representation of the HER2 and basal subtypes in BRCA-associated carcinomas. *BMC Cancer* 2012; **12**(1): 1-3.
60. Silvestri V, Rizzolo P, Falchetti M, et al. Mutation analysis of BRIP1 in male breast cancer cases: a population-based study in Central Italy. *Breast Cancer Res Treat* 2011; **126**(2): 539-43.
61. Ding YC, Steele L, Kuan CJ, Greilac S, Neuhausen SL. Mutations in BRCA2 and PALB2 in male breast cancer cases from the United States. *Breast Cancer Res Treat* 2011; **126**(3): 771-8.
62. Besic N, Cernivc B, de Greve J, et al. BRCA2 gene mutations in Slovenian male breast cancer patients. *Genet Test* 2008; **12**(2): 203-9.
63. Wasielewski M, Den Bakker MA, Van Den Ouweland A, et al. CHEK2 1100delC and male breast cancer in the Netherlands. *Breast Cancer Res Treat* 2009; **116**(2): 397-400.
64. Chodick G, Struewing JP, Ron E, Rutter JL, Iscovich J. Similar prevalence of founder BRCA1 and BRCA2 mutations among Ashkenazi and non-Ashkenazi men with breast cancer: Evidence from 261 cases in Israel, 1976-1999. *Eur J Med Gen* 2008; **51**(2): 141-7.
65. Falchetti M, Lupi R, Rizzolo P, et al. BRCA1/BRCA2 rearrangements and CHEK2 common mutations are infrequent in Italian male breast cancer cases. *Breast Cancer Res Treat* 2008; **110**(1): 161-7.
66. Tchou J, Ward MR, Volpe P, et al. Large genomic rearrangement in BRCA1 and BRCA2 and clinical characteristics of men with breast cancer in the United States. *Clin Breast Cancer* 2007; **7**(8): 627-33.
67. Syrjäkoski K, Fredriksson H, Ikonen T, et al. Hemochromatosis gene mutations among Finnish male breast and prostate cancer patients. *Int J Cancer* 2006; **118**(2): 518-20.
68. Miolo G, Della Puppa L, Santarosa M, et al. Phenotypic features and genetic characterization of male breast cancer families: identification of two recurrent BRCA2 mutations in north-east of Italy. *BMC Cancer* 2006; **6**(1): 1-9.
69. Syrjakoski K, Kuukasjarvi T, Waltering K, et al. BRCA2 mutations in 154 finnish male breast cancer patients. *Neoplasia* 2004; **6**(5): 541-5.
70. Tournier I, Brigitte BDP, Sobol H, et al. Significant contribution of germline BRCA2 rearrangements in male breast cancer families. *Cancer Res* 2004; **64**(22): 8143-7.
71. Ottini L, Masala G, D'Amico C, et al. BRCA1 and BRCA2 mutation status and tumor characteristics in male breast cancer: A population-based study in Italy. *Cancer Res* 2003; **63**(2): 342-7.

72. Gudmundsdottir K, Thorlacius S, Jonasson JG, Sigfusson BF, Tryggvadottir L, Eyfjord JE. CYP17 promoter polymorphism and breast cancer risk in males and females in relation to BRCA2 status. *Br J Cancer* 2003; **88**(6): 933-6.
73. Evans DG, Bulman M, Young K, Gokhale D, Lalloo F. High detection rate for BRCA2 mutations in male breast cancer families from North West England. *Fam Cancer* 2001; **1**(3-4): 131-3.
74. Basham VM, Lipscombe JM, Ward JM, et al. BRCA1 and BRCA2 mutations in a population-based study of male breast cancer. *Breast Cancer Res* 2002; **4**(1): 1-5.
75. Wolpert N, Warner E, Seminsky MF, Futreal A, Narod SA. Prevalence of BRCA1 and BRCA2 mutations in male breast cancer patients in Canada. *Clin Breast Cancer* 2000; **1**(1): 57-63.
76. Díez O, Cortés J, Domènech M, et al. BRCA2 germ-line mutations in Spanish male breast cancer patients. *Ann Oncol* 2000; **11**(1): 81-4.
77. Csokay B, Udvarhelyi N, Sulyok Z, et al. High frequency of germ-line BRCA2 mutations among Hungarian male breast cancer patients without family history. *Cancer Res* 1999; **59**(5): 995-8.
78. Haraldsson K, Loman N, Zhang QX, Johannsson O, Olsson H, Borg A. BRCA2 germ-line mutations are frequent in male breast cancer patients without a family history of the disease. *Cancer Res* 1998; **58**(7): 1367-71.
79. Mavraki E, Gray IC, Bishop DT, Spurr NK. Germline BRCA2 mutations in men with breast cancer. *Br J Cancer* 1997; **76**(11): 1428-31.
80. Friedman LS, Gayther SA, Kurosaki T, et al. Mutation analysis of BRCA1 and BRCA2 in a male breast cancer population. *Am J Hum Genet* 1997; **60**(2): 313-9.
81. Thorlacius S, Sigurdsson S, Bjarnadottir H, et al. Study of a single BRCA2 mutation with high carrier frequency in a small population. *Am J Hum Genet* 1997; **60**(5): 1079-84.
82. Anelli A, Anelli TFM, Youngson B, Rosen PP, Borgen PI. Mutations of the p53 gene in male breast cancer. *Cancer* 1995; **75**(9): 2233-8.
83. Deb S, Gorringe KL, Pang JMB, et al. BRCA2 carriers with male breast cancer show elevated tumour methylation. *BMC Cancer* 2017; **17**(1): 1-11.
84. Kornegoor R, Moelans CB, Verschuur-Maes AHJ, et al. Oncogene amplification in male breast cancer: analysis by multiplex ligation-dependent probe amplification. *Breast Cancer Res Treat* 2012; **135**(1): 49-58.
85. Zanna I, Silvestri V, Palli D, et al. Smoking and FGFR2 rs2981582 variant independently modulate male breast cancer survival: A population-based study in Tuscany, Italy. *Breast* 2018; **40**: 85-91.

86. Chen L, Weng YM, Hu MX, Peng M, Song QB. Effects of HER2 status on the prognosis of male breast cancer: A population-based study. *Onco Targets Ther* 2019; **12**: 7251-60.
87. André S, Pereira T, Silva F, et al. Male breast cancer: Specific biological characteristics and survival in a Portuguese cohort. *Mol Clin Oncol* 2019; **10**(6): 644-54.
88. Moelans CB, De Ligt J, Van Der Groep P, et al. The molecular genetic make-up of male breast cancer. *Endocr Relat Cancer* 2019; **26**(10): 779-94.
89. Gargiulo P, Pensabene M, Milano M, et al. Long-term survival and BRCA status in male breast cancer: A retrospective single-center analysis. *BMC Cancer* 2016; **16**(1): 1-11.
90. Piscuoglio S, Ng CK, Murray MP, et al. The Genomic Landscape of Male Breast Cancers. *Clin Cancer Res* 2016; **22**(16): 4045-56.
91. Abreu MH, Gomes M, Menezes F, et al. CYP2D6 4 polymorphism: A new marker of response to hormonotherapy in male breast cancer? *Breast* 2015; **24**(4): 481-6.
92. Johansson I, Ringner M, Hedenfalk I. The Landscape of Candidate Driver Genes Differs between Male and Female Breast Cancer. *PLoS One* 2013; **8**(10): e78299.
93. Kornegoor R, Verschuur-Maes AHJ, Buerger H, et al. Immunophenotyping of male breast cancer. *Histopathology* 2012; **61**(6): 1145-55.
94. Song YN, Geng JS, Liu T, et al. Long CAG Repeat Sequence and Protein Expression of Androgen Receptor Considered as Prognostic Indicators in Male Breast Carcinoma. *PLoS One* 2012; **7**(12): e52271.
95. Fonseca RR, Tomas AR, Andre S, Soares J. Evaluation of ERBB2 gene status and chromosome 17 anomalies in male breast cancer. *Am J Surg Pathol* 2006; **30**(10): 1292-8.
96. Kwiatkowska E, Teresiak M, Filas V, Karczewska A, Brębborowicz D, Mackiewicz A. BRCA2 mutations and androgen receptor expression as independent predictors of outcome of male breast cancer patients. *Clin Cancer Res* 2003; **9**(12): 4452-9.
97. Zhang S, Liu B, Zhou M, et al. The beneficial role of Asian-based RecurIndex test in the prognostic prediction in Chinese male breast cancer patients. *Sci Rep* 2021; **11**(1): 1-8.
98. Severson TM, Kim Y, Joosten SEP, et al. Characterizing steroid hormone receptor chromatin binding landscapes in male and female breast cancer. *Nat Commun* 2018; **9**(1): 1-12.
99. Massarweh SA, Sledge GW, Miller DP, McCullough D, Petkov VI, Shak S. Molecular Characterization and Mortality From Breast Cancer in Men. *J Clin Oncol* 2018; **36**(14): 1396-404.
100. André S, Pinto AE, Silva GL, Silva F, Serpa J, Félix A. Male Breast Cancer-Immunohistochemical Patterns and Clinical Relevance of FASN, ATF3, and Collagen IV. *Breast Cancer: Basic Clin Res* 2021; **15**: 1-14.

Figure legend

Figure 1

(A) MBC biomarkers that were investigated across multiple omics categories aligned to their associated survival outcomes if present; (B) MBC biomarkers that had associations with multiple hallmarks of cancer aligned to their associated survival outcomes if present.

Table 1: (A) common proteomic biomarkers in breast cancer, **(B)** other well-studied proteomic biomarkers in MBC and their effects on prognosis

Protein biomarkers	**Effects on prognosis**
(A) Common biomarkers	
ERα	***Positivity predicts:*** Improved OS* (frequency = 99.3%[7], 87.6%[104], and 32%[134]; all p < 0.05)[7,104,134]; improved DFS* (frequency = 99.3%; p = 0.001)[7]; improved DSS* (frequency = 93%; p < 0.01)[121] ***Positivity associated with:*** Low Ki-67 index (frequency = 93.1%[87] and 91%[133]; both p < 0.05)[87,133]; PR positivity (frequency = 82%; p = 0.01)[202]; AR positivity (frequency = 91%; p = 0.036)[133]; Bcl-2 positivity (frequency = 82%; p = 0.04)[202]; pS2 positivity (frequency = 82%; p = 0.04)[202]; >60 years of age at diagnosis (frequency = 82%; p = 0.03)[202] ***More frequently expressed in:*** MBCs* compared to FBCs* in general (frequency = 100% vs 86%[136] and 82.3% vs 53.4%[185]; both p < 0.05)[136,185]; MBCs compared to post-menopausal FBCs* (frequency = 82.3% vs 48.9%; p = 0.01)[185] ***Other:*** Lower intensity of expression in MBCs* compared to FBCs* of age group 26-35 years (p = 0.001)[191]; higher median tumour levels in MBCs* compared to FBCs* (p = 0.02)[135]
PR	***Positivity predicts:*** Improved OS* (frequency = 81.9%[7]; 67.2%[104], and 80%[105]; all p < 0.05)[7,104,105]; improved DFS* (frequency = 81.9%; p = 0.002)[7]; improved DSS* (frequency = 77%; p = 0.01)[121]; reduced OS* (p = 0.036)[103]**; reduced DFS* (p = 0.01)[103]** ***Positivity associated with:*** Low Ki-67 index (p < 0.001); low pathological stage (p = 0.029); *BRCA2* mutation negativity (p = 0.01). Frequency = 75.2%[87] ***Other:*** Higher frequency of positivity in MBCs* compared to FBCs* (frequency = 91% vs 76%[136] and 77% vs 62%[202]; p = 0.01)[136,202]; lower intensity of expression in MBCs* compared to FBCs* of age group 26-35 years (p = 0.001)[191]; higher median tumour levels in MBCs* compared to FBCs* (p = 0.04)[135]
ERα/PR co-expression	***Positivity predicts:*** Improved OS* (frequency = 78.1%; p = 0.0054)[118]; improved DFS* (p = 0.022)[118] ***Positivity associated with:*** Low Ki-67 index (frequency = 78.1%; p = 0.029)[118]

HER2	***Positivity predicts:*** Reduced OS* (frequency = 8%[95], 13.5%[101], and 56%[129]; all p < 0.05)[95,101,129]; reduced OS* in ERα positive cases (p = 0.003)[6]; reduced DSS* (p = 0.0001)[101] ***Positivity associated with:*** Younger age of diagnosis (frequency = 13.5%; p < 0.001)[101]; large tumour size (frequency = 3%; p < 0.001)[188]; distant metastasis (frequency = 11%; p = 0.009)[87]; high Ki-67 index (frequency = 11%; p = 0.011)[87]; high anatomic stage (frequency = 11%; p = 0.015)[87]; high tumour grade (frequency = 3%[188] and 62.5%[198]; both p < 0.05)[188,198]
AR	***Positivity predicts:*** Improved OS* in general (frequency = 96.9%[7] and 62.5%[116]; both p < 0.05)[7,116]; improved DFS* in general (frequency = 96.9%[7]; both p < 0.05)[6,7]**; improved 5-year OS* in Luminal A MBCs* compared to Luminal A FBCs* (frequency = 64%; p = 0.01)[123]; reduced 5-year OS* in general (frequency = 82.7%[94], 55.8%[96], and 40.2%[117]; all p < 0.05)[94,96,117]; reduced DFS* in general (frequency = 55.8%; p = 0.002)[96]; reduced 5-year DFS* (frequency = 82.7%[94] and 40.2%[117]; both p < 0.05)[94,117] ***Positivity associated with:*** ERα positivity (frequency = 82.7%[94], 62.5%[116], and 34%[131]; all p < 0.05)[94,116,131,179]**; PR positivity (frequency = 82.7%; p = 0.024)[94]; older age at diagnosis (frequency = 38.5%; p = 0.05)[200]; low proliferative activity (frequency = 34%; p = 0.04)[131]; low tumour grade (p < 0.05)[179]**; poor clinical benefit (frequency = 40.2%; p = 0.025)[117]; node positivity (frequency = 40.2%; p = 0.032)[117]; node negativity in cases with <20% PR positivity (p = 0.007)[179]** ***Other:*** Higher frequency of positivity in MBCs* compared to FBCs* (frequency = 94% vs 63%; p < 0.0001)[123]
Ki-67/MIB1	***High Ki-67 / MIB-1 index predicts:*** Reduced OS* (frequency = 58.9%[87], 48%[129], 46.8%[131], and 48.2%[135]; all p < 0.05)[87,129,131,135]; reduced DFS* (frequency = 58.9%; p = 0.03)[87]; reduced PFS* (frequency = 38%; p = 0.012)[133] ***High Ki-67 / MIB-1 index associated with:*** High tumour grade (frequency = 58.9%[87] and 46.9%[118]; all p < 0.05)[87,118,186,196]**; high anatomic stage (frequency = 58.9%; p = 0.004)[87]; node positivity (frequency = 58.9%[87] and 19.4%[197]; both p < 0.01)[87,197]; positive family history (frequency = 58.9%; p = 0.002)[87]; *BRCA2* mutation positivity (frequency = 58.9%; p = 0.047)[87]; ERα/PR co-expression (both p < 0.05)[186,200]**
(B) Other biomarkers	**Effects on prognosis**
p53	***Positivity predicts:*** Reduced 10-year OS (frequency = 21.2%; p = 0.015)[119] ***Positivity associated with:*** ERα negativity (frequency = 13.6%; p = 0.002)[202]; PR negativity (frequency = 13.6%; p < 0.001)[202]; Bcl-2 negativity (frequency = 13.6%; p = 0.02)[202]; node metastases (frequency = 15%[93] and 16.7%[181]; both p < 0.05)[93,181]; tumour grade 3 (overexpression) (frequency = 15%; p = 0.049)[93]

	Other: Positivity[128,129,131] / overexpression[93] independently predicts reduced OS (frequency = 54%[128], 54%[129], 57.4%[131], and 15%[93]; all $p < 0.05$); negativity associated with Luminal A type (frequency = 78.8%[119] and 83.3%[181]; both $p < 0.05$)[119,181]; higher frequency of positivity in FBCs compared to MBCs (frequency = 18% vs 4%; $p < 0.001$)[160]
Bcl-2	***Positivity associated with:*** ERα positivity (frequency = 94%; $p = 0.04$)[189]; PR positivity (frequency = 56.6%; $p = 0.008$)[194]; node positivity (frequency = 66.7%[181] and 56.6%[194]; both $p < 0.05$)[181,194]; small tumour size (frequency = 73%; $p = 0.017$)[93] ***Negativity associated with***: Luminal B type ($p = 0.028$); tumour grade 3 ($p = 0.01$), frequency = 25%[93] ***Other:*** Higher frequency of positivity in MBCs* compared to FBCs* (frequency = 67% vs 48%; $p = 0.006$)[202]
Cyclin D1	***Positivity predicts:*** Improved PFS* (frequency = 58%; $p = 0.009$)[133]; improved DFS* (frequency = 83.7%; $p = 0.04$)[125]; improved DSS* ($p = 0.001$)[121]** ***Positivity associated with:*** Small tumour size (frequency = 77%[93] and 83.7%[125]; both $p < 0.05$)[93,125]; node negativity (frequency = 83.7%; $p = 0.04$)[125]; p53 positivity (frequency = 58%; $p < 0.001$)[133]; AR positivity (frequency = 58%; $p = 0.028$)[133]
Hypoxic biomarkers **HIF1-α**	***Positivity predicts:*** Reduced DSS* in sporadic MBCs* but not familial MBCs* (frequency = 59% vs 15.5%; $p = 0.006$)[141]; overexpression independently predicts reduced DSS* (frequency = 27%; $p < 0.05$)[124]; perinecrotic staining predicts reduced OS* (frequency = 22.4%; $p = 0.014$)[124†]; diffuse staining in >5% tumour cells associated with high histological grade ($p < 0.001$) and high mitotic count ($p = 0.038$; frequency = 34.4%)[124] ***Positivity associated with***: Invasive carcinoma of no special type ($p = 0.005$); basal cell intrinsic phenotype ($p = 0.02$; frequency = 25.1%)[141] ***Overexpression associated with:*** High tumour grade (frequency = 27%[124] and 36.2%[180]; both $p < 0.05$)[124,180]; high mitotic activity (frequency = 36.2%; $p = 0.013$)[180]; HER2 amplification (frequency = 27%; $p = 0.005$)[124]; Glut-1 overexpression (frequency = 27%; $p < 0.001$)[124]; CA-9 overexpression (frequency = 27%; $p = 0.034$)[124] ***Other:*** High similarity of expression between invasive carcinoma and adjacent DCIS* (frequency = 36.2% vs 37.9%; $p < 0.001$)[180]; higher frequency of Glut-1/CA-9 overexpression with HIF1-α perinecrotic staining compared to diffuse staining in DCIS* (both pure and adjacent) (frequency = 60% vs 100%; $p = 0.012$)[180];

CA-9	***Positive expression predicts:*** Reduced DSS* (frequency = 8%; p = 0.002)[141] ***Other:*** High similarity of expression between invasive carcinoma and adjacent DCIS* (frequency = 37.9% vs 24.1%; p < 0.001)[180]
HIF1-α and/or CA-9 expression	***Expression of either marker predicts:*** Reduced DSS* (frequency = 25.1% and 8% for HIF1-α and CA-9 respectively; p = 0.008)[141]
Glut-1	***Overexpression associated with:*** High mitotic count (p = 0.014); high tumour grade (p = 0.038; frequency = 62.1% for invasive carcinoma[180] ***Other:*** High similarity of expression between invasive carcinoma and adjacent DCIS* (frequency = 75.8% vs 62.1%; p < 0.001)[180]
p21	***Positivity predicts:*** Reduced DFS* (frequency = 41.3%; p = 0.04)[125] ***Positivity associated with***: HER2 negativity (frequency = 70.3%; p = 0.05)[196]; high mitotic activity (frequency = 48%; p < 0.001)[93]; tumour grade 3 (frequency = 48%; p = 0.002)[93] ; Luminal B type (frequency = 48%; p = 0.026)[93] ***Other:*** Higher frequency of positivity in MBCs* compared to FBCs* (frequency = 96% vs 58%[160] and 70.3% vs 29%[196]; both p < 0.01)[160,196]
p27	***Negativity associated with:*** Lymph node metastases (frequency = 81.2%[125] and 64%[197]; both p < 0.05)[125,197] ***Overexpression associated with:*** AR positivity (frequency = 96.2%; p = 0.049)[196] ***Other:*** Higher frequency of positivity in MBCs* compared to FBCs* (frequency = 96.2% vs 39.3%; p = 0.00)[196]
EGFR	***Overexpression associated with***: HER2 amplification (frequency = 12%; p = 0.04)[190] ***Positivity associated with:*** ERα and PR negativity (frequency = 11.4%; both p = 0.04)[188]; high MIB-1 index (frequency = 9.4%; p = 0.0181)[118]
c-Myc	***Positivity predicts:*** Reduced OS* (frequency = 82%; p = 0.01)[129]

	Other: Overexpression predicts improved DFS* (frequency = 90%; p = 0.04)[125] and is associated with node negativity (frequency = 90%; p = 0.006)[125]

**MBC: Male Breast Cancer; FBC: Female Breast Cancer; OS: Overall Survival; DFS: Disease Free Survival; DSS: Disease Specific Survival; PFS: Progression Free Survival; DCIS: Ductal Carcinoma In-Situ*

***frequency unavailable from all/some source article(s)*

†*Perinecrotic staining: Staining surrounding a necrotic area*

Table 2: Ten most studied genetic/transcriptomic biomarkers in MBC and their effects on prognosis

Biomarker	Effects on prognosis
BRCA2	***Mutated status predicts:*** Reduced OS* in general (frequency = 10.8%[85] and 29.5%[87]; both $p < 0.05$)[85,87]; reduced 5-year OS* (frequency = 27.9%; $p = 0.003$)[96]; reduced DSS* in general (frequency = 29.5%; $p = 0.003$)[87]; reduced 5-year DSS* (frequency = 27.9%; $p = 0.006$)[96] ***Mutated status associated with:*** ERα negativity (frequency = 9.3%; $p = 0.05$)[173]; PR negativity (frequency = 29.5%[87], 12.2%[170] and 9.3%[173]; all $p < 0.05$)[87,170,173]; HER2 positivity/enriched subtype (frequency = 12.2%[170] and 9.3%[173]; both $p < 0.05$)[170,173]; Luminal B type (frequency = 12.2%; $p = 0.016$)[170]; advanced tumour grade[164,173]/ tumour grade 3[61,170] (frequency = 89.4%[164†], 9.3%[173], 15.6%[61], and 12.2%[170]; all $p < 0.05$); higher frequency of tumour grade 3 in patients <50 years of age (frequency = 89.4%; $p = 0.005$)[164†]; node positivity (frequency = 15.6%; $p < 0.02$)[61]; contralaterality (frequency = 12.2%; $p = 0.01$)[170]; bilaterality (frequency = 29.5%; $p = 0.008$)[87]; high Ki-67 index (frequency = 29.5%; $p = 0.047$)[87]; higher frequency of genetic aberrations in *BRCA2*-mutated MBCs compared to *BRCA2*-wt MBCs ($p < 0.05$)[175**]; family history of breast/ovarian cancer or personal history of cancer (frequency = 12.2%[170]; all $p < 0.05$)[68,170**]; amplification of *CCNE2, ASAP1, CSMD3, UBR5, DNAH11, RRM2B, FZD6, RUNX1T1* and *SGK3* (frequency = 11%; all $p < 0.05$)[88]; decreased copy number aberration load on chr 8p (frequency = 11%; $p = 0.004$)[88] ***Other:*** Higher frequency of mutations in MBCs* compared to FBCs* (frequency = 41.7% vs 8.3%; $p = 0.0008$)[59]; higher tumour grade in *BRCA2*-mutated MBCs* compared to SEER* MBCs* ($p = 4.52e\text{-}12$)[164]; higher disease stage in *BRCA2*-mutated MBCs* compared to *BRCA2*-mutated FBCs* ($p = 2.14e\text{-}5$)[164]; increased disease risk in men <60 years (OR* = 5.63; frequency = 29.4%; $p < 0.05$)[149]
HER2	***Amplified status predicts****:* Reduced OS* in general[86,95] – also predicted by copy number gain[84] (frequency = 13.3%[86], 8%[95], and 4%[84]; all $p < 0.05$); reduced 4-year OS* (frequency = 13.3%; $p = 0.005$)[86]; reduced OS* in patients with tumour size of 2-4 cm (frequency = 13.3%; $p = 0.02$)[86]; reduced OS* in patients with distant metastasis (frequency = 13.3%; $p = 0.023$)[86]; reduced OS* in patients who have undergone radiation therapy (frequency = 13.3%; $p = 0.041$)[86] ***Amplified status associated with****:* High mean mitotic activity (frequency = 3%; $p < 0.001$)[93]; poor degree of differentiation[86] / histological grade 3[93] (frequency = 13.3%[86] and 3%[93]; both $p < 0.05$); distant metastasis (frequency = 13.3%; $p = 0.002$)[86]; regional lymph node metastasis (frequency = 13.3%; $p = 0.004$)[86]; younger age of diagnosis (frequency = 13.3%; $p < 0.001$)[86]; large tumour size (frequency = 13.3%; $p < 0.001$)[86]; advanced disease stage (frequency = 13.3%; $p < 0.001$)[86]; surgery and chemotherapeutic treatment (frequency = 13.3%; $p < 0.001$)[86] ***Other:*** Downregulated in MBCs* compared to FBCs* ($p < 0.01$)[171**]

CCND1	***Amplified status associated with:*** ERα positivity (frequency = 63%; p < 0.0001)[174]; HER2 positivity (frequency = 16%; p = 0.0005)[165]; high MIB-1 index (frequency = 16%; p = 0.04)[165] ***Amplified status predicts***: Reduced OS* (frequency = 46%; p = 0.022)[84] ***Other:*** Higher copy number ratio and amplification frequency in high grade invasive carcinoma compared to low/intermediate grade invasive carcinoma (all p = 0.005)[162]**
PALB2	***Associations with MBC risk:*** Pathogenic variants associated with MBC risk (control dataset specific results; frequency = 1.2%)[54]; EVS* dataset: OR = 17.30 (p < 0.0001); ExAc* dataset: OR = 11.20 (p < 0.0001); gnomAD* dataset: OR = 9.63 (p < 0.0001) ***Other:*** Copy number gain (exon 6) associated with node negativity (p = 0.021)[12]**; Mutated status associated with bilaterality (frequency = 2.4%; p = 0.004)[49]; Higher frequency of mutations in MBC* compared to unmatched female normal breast tissue (frequency = 2.4%; p < 0.001)[49]
PIK3CA	***Mutated status associated with:*** *BRCA2* mutation negativity (frequency = 10.5%; p = 0.03)[169]; node positivity (frequency = 36.1%; p = 0.006)[88]; advanced tumour grade (frequency = 36.1%; p = 0.013)[88]; high mitotic index (frequency = 36.1%; p = 0.014)[88]; absence of both nuclear and cytoplasmic expression of p4E-BP1 (frequency = 10.5%; both p < 0.05)[169]; pS6 upregulation (frequency = 10.5%; p = 0.024)[169] ***Less frequently mutated in:*** ERα positive/HER2 negative MBCs* compared to matched FBCs* (frequency = 18% vs 42%; p = 0.0005)[90]; ERα positive/HER2 negative MBCs* compared to matched post-menopausal FBCs* (frequency = 18% vs 42%; p = 0.0014)[90]
GATA3	***Mutated status***: predicts reduced DFS* (frequency = 15%; p = 0.038)[90]; associated with Luminal B type (frequency = 15%; p = 0.0482)[90] ***Other:*** Upregulation associated with AR positivity (p = 0.0347)[171]**
EGFR	***Amplification associated with:*** ERα negativity (p = 0.01); HER2 positivity (p = 0.03); stage IV disease (p = 0.01). Amplification frequency = 6.8%[165] ***Other:*** Copy number gain associated with high grade invasive carcinoma (frequency = 62%; p = 0.047)[162]††
EMSY	***Amplification predicts:*** Reduced OS* (p = 0.04)[84]** ***Amplification associated with:*** *BRCA1/2* mutation positivity (frequency = 34.7%; p = 0.03)[163]

miR-125b	***High expression:*** associated with small tumour size (p = 0.03)[209]** ***Downregulated:*** MBCs* compared to FBCs* (p < 0.01); MBCs* compared to gynaecomastia (p < 0.01)[177]
rs3803662 (*TOX3;* risk biomarker)	**Associated with MBC* risk:** OR* = 1.48 (p = 4e-6)[145]**; OR* = 1.59 (frequency = 34.7%, 47.3%, and 18% for CC, CT, and TT genotypes, respectively; p = 0.0001)[144]

**MBC: Male Breast Cancer; FBC: Female Breast Cancer; OS: Overall Survival; DFS: Disease Free Survival; DSS: Disease Specific Survival; SEER: Surveillance Epidemiology and End Results; EVS: Exome Variant Server; ExAC: Exome Aggregation Consortium; gnomAD: Genome Aggregation Database*

***Breakdown for gene-specific alteration unavailable from all or some source articles*

†Cohort selected for BRCA1/2 mutations

††Frequency of CNV in pure ductal carcinoma in-situ (DCIS): 6% (CCND1 amplification), 6% (EGFR gain) and in DCIS adjacent to invasive carcinoma (DCIS-AIC): 16% (CCND1 amplification), 2% (EGFR gain)

Table 3: Ten most studied epigenetic biomarkers in MBC and their effects on prognosis

Biomarker	Effects on prognosis
ESR1	***Promoter hypermethylation:*** Associated with high tumour grade (p = 0.037); high mean mitotic count (p = 0.001), frequency = 8%[15] ***Other:*** Promoter hypermethylation less frequent in MBC* compared FBC* (frequency = 8%; p = 0.005)[15]; higher methylation in tumours compared to peripheral blood (p < 0.0001)[156]**; lower absolute methylation % in male DCIS-AIC* compared to female DCIS-AIC* (frequency of hypermethylated cases† in male DCIS-AIC = 5%; p < 0.002)[155]
GSTP1	***Promoter hypermethylation:*** Associated with high tumour grade (frequency = 44%; p = 0.001)[15]; high mean mitotic count (frequency = 44%; p = 0.002)[15]; *BRCA2* mutation positivity (frequency = 82%; p = 0.02)[83] ***Other:*** High absolute methylation % associated with high grade invasive carcinoma (frequency = 41%; p = 0.047)[155]
RARB	***Promoter hypermethylation:*** Associated with ERα negativity (frequency = 8%; p = 0.04)[157]; PR positivity (frequency = 8%; p = 0.03)[157]; large tumour size (frequency = 30%; p = 0.01)[83]; presence of Paget's disease (frequency = 30%; p = 0.01)[83]; *BRCA2* mutation positivity (frequency = 30%; p = 0.02)[83]; less frequent in MBC* compared FBC* (frequency = 5% vs 20%; p = 0.026)[15]
RASSF1/RASSF1A	***Promoter hypermethylation:*** Associated with ERα negativity (frequency = 76%; p = 0.0001)[157] ; PR positivity (frequency = 76%; p = 0.00)[157] ; HER2 positivity (frequency = 79.1%; p = 0.01)[156]; presence of DCIS* (frequency = 68%; p = 0.02)[83]; *BRCA1/2* mutation positivity (frequency = 79.1%; p = 0.008)[156]; tumour grade G3 (frequency = 79.1%; p = 0.008)[156]; more frequent in MBC* compared to FBC* (frequency = 76% vs 28%; p = 0.0001)[157] ***Other:*** Higher methylation levels in tumours compared to peripheral blood (p < 0.0001)[156]
AR	***Promoter hypermethylation:*** Associated with *BRCA1/2* mutation negativity (frequency = 94%; p = 0.016)[156] ***Other:*** CpG hypomethylation in MBC* cases compared to gynaecomastia cases (p < 0.05)[154.] Higher methylation in tumours compared to male normal breast tissue (p = 0.0009); tumours compared to lymph nodes (p = 0.003); tumours compared to peripheral blood (p = 0.0006). Frequency = 94%[156]
ATM	***Promoter hypermethylation:*** Less frequent in MBC* compared FBC* (frequency = 1% vs 15%; p = 0.017)[15] ***Other:*** High absolute methylation % associated with high grade invasive carcinoma (p = 0.036)[155]††

BRCA2	***Promoter hypermethylation:*** Less frequent in MBC* compared FBC* (frequency = 17% vs 60%; p < 0.001)[15] ***Other:*** Lower absolute methylation % in male DCIS-AIC* compared to female DCIS-AIC* (p < 0.02)[155]
MGMT	***Promoter hypermethylation:*** Associated with larger mean tumour size than tumours without *MGMT* hypermethylation (frequency = 7%; p = 0.002)[15]; higher frequency in pure invasive carcinoma compared to DCIS-AIC* (frequency = 25% vs 9%; p = 0.039)[155]
VHL	***Promoter hypermethylation:*** Less frequent in MBC* compared to FBC* (frequency = 2% vs 15%; p = 0.025)[15] ***Other:*** Lower absolute methylation % in male DCIS-AIC* compared to female DCIS-AIC* (p < 0.002)[155††]
TWIST1	***Promoter hypermethylation predicts:*** Reduced DSS* in *BRCA2* mutation positive MBC patients (p = 0.001); reduced DSS* in all MBC patients (p = 0.01). Frequency = 37%[83]

**MBC: Male Breast Cancer; FBC: Female Breast Cancer; DSS: Disease Specific Survival; DCIS: Ductal Carcinoma In-Situ; DCIS-AIC: Ductal Carcinoma In-situ Adjacent to Invasive Carcinoma*

***Frequency unavailable from source article*

†Frequency of ESR1 hypermethylated cases in male pure-DCIS = 6% and invasive carcinoma = 9%; frequency of BRCA2 hypermethylated cases in male pure-DCIS = 11% and invasive carcinoma = 2%

††Promoter hypermethylation was not present in the MBC cohort. However, higher absolute methylation % of ATM was observed in high grade tumours compared to low/intermediate grade tumours. Similarly, lower absolute methylation % of VHL was observed in male DCIS-AIC compared to female DCIS-AIC

Table 4: Ten most studied morphological features in MBC and their effects on prognosis

Morphological feature	Effects on prognosis
TIL* density	***High density of TILs*:*** Predicts improved OS* (p = 0.011) and RFS* (p =0.02, frequency = 14.3%)[137]; association with node positivity (frequency = 27.8%; p = 0.025)[186] ***Other:*** Higher density of TILs* in HER2 positive MBCs* compared to Luminal HER2 negative MBCs* (overall frequency of high TIL* density = 14.3%; p = 0.015)[137††]

Fibrotic focus	***Presence of fibrotic foci:*** Predicts reduced OS* (p = 0.004) and RFS* (p < 0.001) at a frequency of 32.2%)[137]; reduced overall survival when foci of >8 mm† (p = 0.035)[124] and associated with (frequency = 25%)[124]; high tumour grade (p = 0.005); few/no tubule formation (p = 0.03); high nuclear grade (p = 0.038); node positivity (p = 0.037)
Mitotic activity index	***High mitotic activity index:*** Predicts reduced OS* (frequency = 32.5%[138]; both p < 0.05)[137,138]**; reduced RFS* (p = 0.024)[137]**
Mean nuclear area	***High mean nuclear area:*** Predicts reduced OS* (frequency = 50%[128] and 32.5%[138]; both p < 0.05)[128,138]; associated with nuclear atypia (frequency = 32.5%; p = 0.032)[138]; aneuploidy (frequency = 50%; p = 0.01)[128]; high mitotic activity index (frequency = 32.5%; p = 0.011)[138]; high MIB-1 index (frequency = 50%; p = 0.02)[128]; high pathological stage (frequency = 50%; p = 0.01)[128]; high tumour grade (frequency = 50%[128] and 32.5%[138]; both p < 0.05)[128,138]; high PCNA* score (frequency = 50%; p = 0.002)[128]; high AgNOR* quantity (frequency = 50%; p < 0.001)[128]
Standard deviation of nuclear area	***High standard deviation of nuclear area:*** Predicts reduced OS* (frequency = 50%; p = 0.02)[128] and is associated with aneuploidy (frequency = 50%; p = 0.001)[128]; high mitotic activity index (frequency = 32.5%; p = 0.014)[138]; high MIB-1 index (frequency = 50%; p = 0.001)[128]; high tumour grade (frequency = 50%[128] and 32.5%[138]; both p < 0.05)[128,138]; high PCNA* score (frequency = 50%; p < 0.001)[128]; high AgNOR* quantity (frequency = 50%; p < 0.001)[128]; p53 positivity (frequency = 50%; p = 0.005)[128]; Bcl-2 negativity (frequency = 50%; p = 0.04)[128]
Mean nuclear perimeter	***High mean nuclear perimeter:*** Predicts reduced OS* (frequency = 50%; p = 0.01)[128] and is associated with aneuploidy (p = 0.005); high MIB-1 index (p = 0.01); high pathological stage (p = 0.03); high tumour grade (p = 0.002); high PCNA* score (p = 0.001); high AgNOR* quantity (p < 0.001), all at 50% frequency[128]
Standard deviation of nuclear perimeter	***High standard deviation of nuclear perimeter:*** Predicts reduced OS* (frequency = 50%; p = 0.009)[128] and is associated with; aneuploidy (p = 0.001); high MIB-1 index (p = 0.003); high pathological stage (p = 0.001); high tumour grade (p = 0.002); high PCNA* score (p = 0.002) ; high AgNOR* quantity (p < 0.001), all at 50% frequency[128]
Nuclear shape factor (Defined as: (4*π*area)/Perimeter2)	***High shape factor:*** Predicts improved OS* (frequency = 42%; both p < 0.05)[128] and is associated with diploidy (p = 0.0007); low MIB-1 index (p = 0.001); low tumour grade (p = 0.0007); p53 negativity (p = 0.005); c-Myc negativity (p = 0.05); low AgNOR* quantity (p = 0.005), all at 42% frequency[128]
Vascular invasion	***High frequency of vascular invasion:*** Associated with ERα/PR negativity (p = 0.0004); high tumour grade (p = 0.035), both at 20% frequency[186]
Tubule formation	***High tubule formation:*** Predicts improved OS* (frequency = 50.5%; p = 0.035)[138]

**MBC: Male Breast Cancer; OS: Overall Survival; RFS: Relapse Free Survival; PCNA: Proliferating Cell Nuclear Antigen; AgNOR: Argyrophillic Nucleolar Organiser Regions; TILs: Tumour Infiltrating Lymphocytes*

***Frequency unavailable from all/some source article(s)*

†Frequency of fibrotic foci >8mm not available from source article

††Surrogate subtype specific breakdown unavailable

Table 5: Novel clusters identified in MBC. Clinical correlations and/or p-values are specified where available.

Category	Cluster	Outcome
Epigenetic	ME1 Epitype (n = 23)[10]	***Associated with:*** Cyclin A positivity (p = 0.012); high fraction of genome alteration (p = 0.0045); high S-phase fraction (p = 0.035); high mitotic activity (p = 1.5e-5) ; luminal M1 transcriptional subgroup[13] ***Compared to the ME2 epitype, ME1 epitype had***: Lower ERα scores (p = 0.048); higher EZH2 expression (p = 3.3e-7); higher activity of proliferation modules (p = 2.8e-7); more frequent hypermethylation of genes involved in epigenetic gene silencing with H3K27me3 (p = 4.4e-153), transcriptional regulation with HOX genes (p = 1.6e-22), cell adhesion pathways (p = 5.6e-5), WNT signalling (p = 2.8e-4), TGF-β signalling (p < 0.001), focal adhesion (p < 0.005), MAPK signalling (p < 0.005), FGFR ligand binding and activation (p < 0.007)
	ME2 Epitype (n = 24)[10]	***Associated with:*** Luminal M2 transcriptional subgroup (p = 0.011)[13]
	Cluster 1 (n = 20)[83]	***Characterised by:*** Hypermethylation of *GSTP1* and *WIF1*; lower methylation levels of *RASSF1A* compared to *MAL*
	Cluster 2 (n = 19)[83]	***Characterised by***: hypermethylation of *GSTP1*
	Cluster 3 (n = 7)[83]	***Characterised by:*** Lower methylation levels of *WIF1* compared to *RASSF1A*; hypermethylation of *RARB* and *GSTP1* and *a*ssociated with *BRCA2* mutation positivity (p = 0.02)
	Cluster 4 (n = 8)[83]	***Characterised by:*** lower methylation levels of *RASSF1A* compared to *TWIST1*
	BRCA2-mutation positive subgroup: Cluster A (n = 12)[83]	***Characterised by:*** Hypermethylation of *GSTP1* and *MAL;* lower *RASSF1A* methylation compared to Cluster B; younger ages of diagnosis compared to other *BRCA2*-mutation positive patients
	BRCA2-mutation positive subgroup: Cluster B (n = 8)[83]	***Characterised by:*** Hypermethylation of *RASSF1A*
Genetic	Luminal M1 (n = 46)[13]	***Associated with:*** HER2 positivity (p = 0.0057); PLAU expression – invasion and metastasis (p = 1.0e-5); AURKA expression – proliferation (p = 0.026)

	Luminal M2 (n = 20)[13]	***Associated with:*** ESR1 expression & ERα positivity (p = 1.3e-8); STAT1 expression – immune response (p = 6.8e-3)
	Male-simple (n = 11)[14]	***Compared to male-complex group, the male-simple group had:*** Lower fraction of altered genome (p = 0.007); lower S-phase fraction (p = 0.02); smaller tumour size (p = 0.004)
	Male-complex (n = 43)[14]	***Characterised by:*** Similarity with the female Luminal B intrinsic subtype; BRCA2 mutation positivity; whole chromosome arm gains
	Cluster A (n = 78)[12]	***Characterised by:*** Partial and whole arm loss of chromosome 16q; higher copy number gain on chromosome 16p compared to Cluster B; higher frequency of loss of chromosome 16q genes compared to Cluster B
	Cluster B (n = 57)[12]	***Characterised by:*** Higher percentage of copy number gain compared to Cluster A; lower frequency of node positivity compared to Cluster A (p = 0.008) and associated with triple negativity (p = 0.042)
	Cluster A (n = 55)[84]	***Characterised by:*** Low rates of copy number gain and amplification.
	Cluster B (n = 51)[84]	***Characterised by:*** Copy number gain in the genes *CCND1, MTDH, CDC6, ADAM9, TRAF4* and *MYC* and independently predicts reduced overall survival (p = 0.009) and associated with high mitotic index (p < 0.001); tumour grade 3 (p = 0.02); large tumour size (p = 0.036)
Transcriptomic	Cluster 1 (n = 41)[11]	**Predicts:** Reduced OS* (p = 0.043) and associated with AURKA signature (proliferation marker) (p = 0.02); HER2 signalling (p = 0.0003); PLAU signature (invasion and metastasis marker) (p = 0.03); STAT1 signature (immune response marker) (p = 0.005)
	Cluster 2 (n = 22)[11]	***Associated with:*** NAT1 upregulation (p = 0.007); CASP3 signature (apoptosis marker) (p = 0.01)
Proteomic	Cluster A1 (Hormone receptor negative) (n = 21)[93]	***Both A1 and A2 clusters:*** Had reduced 5-year overall survival compared to B1 and B2 clusters (p = 0.011) and characterised by ERα negative cases clustering together with PR and AR negative cases; low protein expression of other markers; intermediate histological grade; associated with large tumour size (p = 0.023)
	Cluster A2 (ERα positive high-grade) (n = 37)[93]	***Both A1 and A2 clusters:*** Had reduced 5-year overall survival compared to B1 and B2 clusters (p = 0.011) and characterised by low PR expression; HER2 amplification; high Ki-67 index; accumulation of p21, p16, and p53; expression of EGFR and CK5/6 and associated with: high tumour grade (p = 0.001); high mitotic activity (p < 0.001); node positivity (p = 0.033)
	Cluster B1 (ERα positive intermediate-grade) (n = 34)[93]	***Characterised by:*** Hormone receptor positivity; Bcl-2 and Cyclin D1 positivity; low Ki-67 index; BRST-2 negativity; node negativity

	Cluster B2 (ERα positive low-grade) (n = 37)[93]	***Characterised by:*** Hormone receptor positivity; Bcl-2 and Cyclin D1 positivity; low Ki-67 index; BRST-2 positivity; node positivity
	c-ERβ cluster[123]**	***Characterised by:*** Cytoplasmic expression of both ERβ1 and ERβ2. Also found in FBC*
	PR cluster[123]**	***Characterised by:*** Both PR-A and PR-B isoform action.
	ERα/ERβ/AR cluster[123]**	***Characterised by:*** Collective action of AR with the ER isoforms α, β1, β2, and β5.

**FBC: Female Breast Cancer; OS: Overall Survival*

***breakdown unavailable*

Figure 1

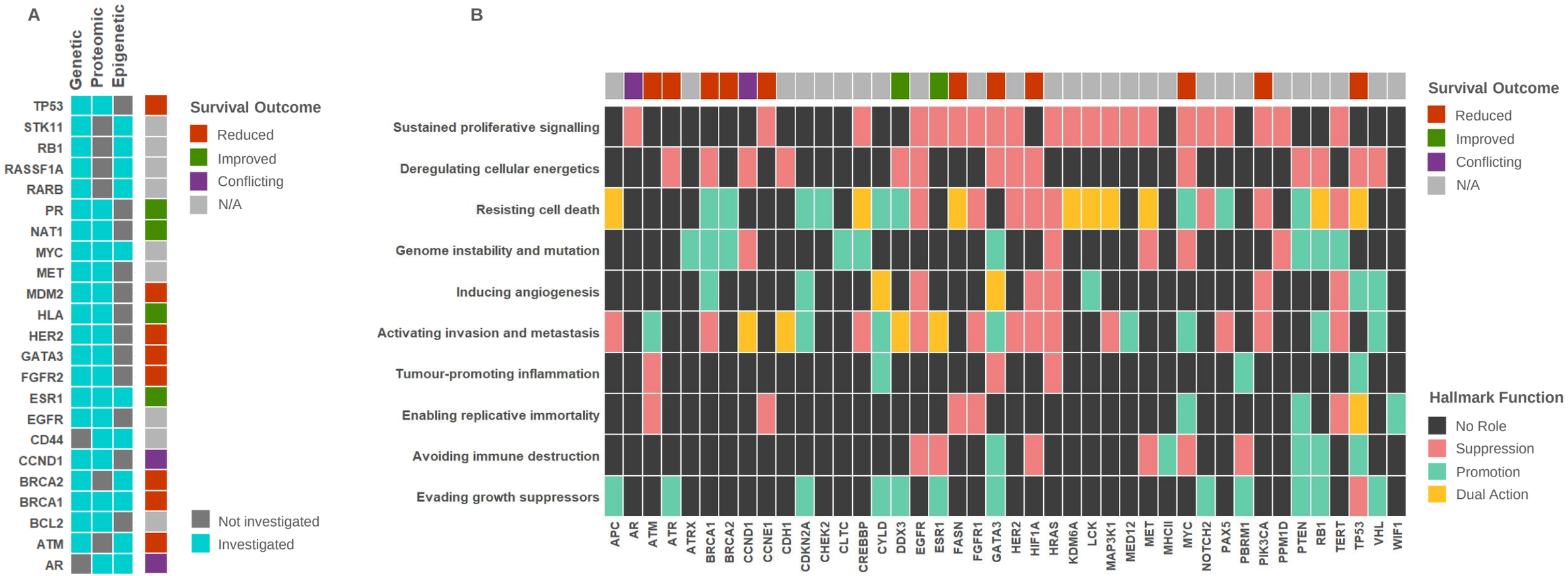

Figure 1: (A) MBC markers that were investigated across multiple omics categories aligned to their survival outcomes (if present); (B) MBC markers that had associations with multiple hallmarks of cancer aligned to their associated survival outcomes (if present).

We thank both reviewers for taking the time to read our article and for their constructive comments. Each of the points raised are discussed below and any changes made to the text have been indicated in red font. Tracks have been removed as several people edited this making the tracked version rather messy to read.

1. **Reviewer #1:**
   Male and female breast cancer differ in their clinical presentation and (epi)genetic makeup but regardless, clinical management of male breast cancer is still largely informed by female-only clinical trials. The authors have performed a systematic review to identify knowledge gaps in the current male breast cancer (MBC) biomarker field. They have comprehensively described a broad spectrum of suggested/potential MBC biomarkers with a focus on prognostic biomarkers, and highlighted several candidates for further investigation. I applaud the authors for their endeavour to systematically combine biomarker data from literature in order to advance MBC research by defining those biomarkers of (potential) translational value. I believe this review is of importance to the field but there are however some points to improve:

   MAJOR comments:

1. The title clearly indicates a focus on prognostic MBC biomarkers. This is however less clear from the abstract and introduction where the term "biomarkers" is often used in general and is very poorly defined. The authors should better define which kind of "biomarkers" they were looking for and why. Also, the BEST working group (https://www.ncbi.nlm.nih.gov/books/NBK326791/) concluded that prognostic biomarkers should be differentiated from susceptibility/risk biomarkers, which deal with association with the transition from healthy state to disease. As the authors also include markers for MBC risk in their review, they should consider making a clear distinction.

   Response: We agree with the recommendations from the BEST working group regarding clarification of biomarkers and have revised the second sentence of the abstract to reflect this more clearly, also at various points in the text e.g. first and last sentences of the last paragraph of introduction, which now refer to prognostic biomarkers and a separate section on pathogenic variations associated with breast cancer risk.

2. In the section "Pathogenic variations in other genes" (starting on page 8, line 188) the authors now mix germline and somatic mutations, which is highly confusing. Please make a clear distinction throughout the review.

   Response: Apologies for the confusion. The 2 subsections are now entitled ‘Pathogenic Variations in BRCA genes with prognostic value’ and ‘Pathogenic variations in other genes with prognostic value’. The mutations mentioned have now been specified as either germline or somatic. Source articles described prognostic associations with only somatic mutations and copy number variations, while risk was only associated with germline mutations, except for BRCA1 and BRCA2. This information has now been specified in the text. Original numbers have changed in light of the text changes.

   The germline/somatic status of all the mutations described in the screening studies have also been specified in Table S6 (Appendix, page 100).

3. I strongly urge the authors to use the terms luminal A/B-like and basal-like when talking about surrogate intrinsic subtypes. Please adjust in text and tables accordingly.

   Response: When we wrote the manuscript, we tried to stay as true as possible to the source articles and one some occasions subtypes were not always clarified. Therefore, the nomenclature throughout the article has been dependent on the reference associated e.g. in line 227-228 of the original paper, the article describing basal MBCs, CK5/6 profiling was conducted but other articles describing triple negativity did not do so. We have changed these where there was no room for doubt from the source articles.

4. The authors list all these potential biomarkers but unfortunately, they do not mention the frequency of their occurrence in the investigated manuscripts (range between studies). This information should be added, especially for the biomarkers that warrant further research and are mentioned individually in the discussion.

   Response: This information has now been added for the markers needing further evaluation and the markers are described in Tables 1-4.

5. Supplementary Table 3 still contains a remark from one of the authors and section B is empty. Please make sure that section B has been added.

   Response: Apologies. This was a formatting error which has now been corrected.

6. As the authors also reference to clinical trials and FDA approvals (for example for AR inhibitors and CDK4/6 inhibitors), I wonder why they did not mention the PIK3CA inhibitor Alpelisib? This biomarker is clearly mentioned in the manuscript and listed in Table 1?

   Response: Thank you for pointing this out. This paper is now discussed in Page 18 Lines 444-449 (tracked manuscript).

MINOR comments and textual changes:

1. It would be interesting to add whether the studied manuscripts described the degree of association between genetics, epigenetics and proteome for each biomarker. For example, the authors mention CCND1/cyclin D1 as an interesting biomarker but with opposing roles for gene alteration and protein alteration. Also, TP53 is almost never mutated in male breast cancer but apparently many studies have investigated its protein overexpression and it is suggested as biomarker that warrants further research (page 16, line 394). Combining (epi)genetic and protein data for these biomarkers could therefore also reveal knowledge gaps.

   Response: This is an excellent suggestion which we now include as a new Figure 1. This 2 panel Figure shows MBC markers that were investigated across multiple omics categories and then aligned to any associated survival outcomes (A). We then present MBC markers that had associations with multiple hallmarks of cancer and which were aligned to any associated survival outcomes (B).

2. On page 5, line 102, manuscript exclusion criteria in the manuscript indicate exclusion if primary cohort size is <5. In Supplementary Fig1 it says <=5. Please clarify.

   Response: This was ≤5 which is now clarified

3. On page 5, line 120, it is unclear to me what the authors mean by "with available clinical features"

   Response: This referred to the clinical features investigated in the source articles and therefore, available to describe in this review. The phrasing has been changed to make this clearer and now reads "… and associations with clinical features described in each article". This is now on p8, line 163 (tracked manuscript).

4. On page 6, quality assessment (line 129), the authors should indicate which checklists were specifically used

   Response: This has been rephrased and now reads "… using Joanna Briggs Institute Critical Appraisal tools using the checklists for case-control studies, and analytical cross-sectional studies, as appropriate[21]". This is now on p8, line 174 (tracked manuscript).

5. On page 9, line 202, please mention whether upregulation or downregulation of miR-125b is associated with survival benefit

   Response: This has been specified to upregulation of miR-125b (p13, line 292, tracked manuscript).

6. On page 10, line 230, I believe that the former should be the latter?

   Response: Apologies. This was a typographical mistake, which has been corrected.

7. Page 11, line 250: "overexpressed in MBC" compared to what? FBC or normal or?

   Response: The comparative cohort was FBC. This is now specified.

8. Page 11, line 272: change lower>more to higher>less (is more logical)

   Response: This has been changed and now reads: "…epigenetic differences, reduced methylation was more common in both invasive carcinoma (IC) and ductal carcinoma in situ…". (p14, line 322)

9. Page 12, line 278-9: hypermethylation lower>less frequent

   Response: This now reads "…was consistently less frequent compared to IC[156]". (p14, line 329, tracked manuscript)

10. Page 14, line 346-8: something is missing in this sentence

Response: This now reads “A bibliometric analysis revealed that most publications in MBC focused on…” (p17, line 403)

11. In the discussion, on page 16, lines 394 and 398, and in Supplementary Figure 3 legend, MDM2 is mentioned twice. Please remove where appropriate.

    Response: Apologies, this is now corrected.

12. Page 16, line 399: remove "in"

    Response: ‘In’ has been replaced by ‘of’

13. Tables in general: make sure that other abbreviations such as DSS are also explained in the legend

    Response: All missing abbreviations have now been added. Apologies for this oversight

14. Table 1, page 47: association with advanced disease (ref 85) is mentioned twice

    Response: The repetition has been removed

15. Table 1, page 49: PIK3CA. As it is associated with BRCA2 mutation negativity, it is not entirely associated with negative prognosis, so perhaps it should say mostly negative, as was done for PR in Table 2?

    Response: Changed to mostly negative

16. Table 2 title is difficult to read. Please make adjustments

    Response: In response to other reviewer comments, we have changed the order of the narrative to make it more logical, meaning Table 2 is now Table 1. The title for this now reads: **“**(A) ERα, PR, HER2, and ERα/PR co-expression profiles and (B) ten most studied additional proteomic markers and their associations with prognosis in MBC

17. Table 5, page 69. Please add the number of cases per subcluster, not only the total amount of patients.

    Response: Added for all articles in Table 5 except Shaaban et al. 2012, where the breakdown was not reported.

Reviewer #2: In the present review, the authors focused on genetic, transcriptomic, proteomic and epigenetic biomarkers as well as phenotypic features with prognostic value in male breast cancer. Overall, the manuscript provides a broad and comprehensive overview of current knowledge in this field, also discussing gaps and limitations of the studies considered.

1. Despite the careful literature research and the considerable amount of studies examined (extensively described in supplementary files), in my opinion, the review should be more focused on prognostic information, highlighting the most relevant and promising markers as well as the most robust associations (e.g. BRCA2-associated MBCs and higher tumor grade in line 180), especially in the sections "Genetic and transcriptomic markers" and "Epigenetic markers".

   Response:

   Thank you for this valuable suggestion. To emphasise the prognostic information, the associations for each marker described in Tables 1-4 have been arranged in the following order: survival outcomes, association with other clinical factors, and difference in expression between comparative groups, i.e., MBC vs FBC, invasive carcinoma vs DCIS etc. Any associations with risk have also been separated, clearly indicated, and put at the end of all other associations.

2. In particular, I would suggest revising the section "Genetic and transcriptomic markers", specifying the difference between germline and somatic alterations and more clearly describing germline alterations not only involved in MBC risk, but also with potential prognostic value.

   Response: This was also raised by the other reviewer. This has been done and is detailed under Major comments, point 2 above.

3. Authors should verify the accuracy of the OR data for PALB2 and RAD51D (line 192) as well as the description of the SUL1A1 polymorphism (line 194).

   Response:

   In the source paper (Rizzolo et al., 2019), OR data for PALB2 and RAD51D was provided compared to 2 separate control cohorts, the EVS (US) and ExAc (European) cohorts. We omitted to include this in the original narrative. This has now been corrected to read "Regarding MBC risk, germline mutations in mutated PALB2 and RAD51D[54] had the highest odds-ratios (OR = 17.30, 8.58; 11.20, 10.18, using the Exome Variant Server and Non-Finnish European datasets, respectively).

   The SULT1A1 polymorphism has been specified as A/A.

4. I would also suggest moving the description of the transcriptomic markers (line 201) to the end of the paragraph.

   Response: This is the final sentence of the paragraph, but we appreciate this might not have been clear as there is as no space between the 2 paragraphs.

5. I would suggest adding an additional and separate section dealing with Hallmarks of Cancer (line 280-284) and moving lines 394-400 to this new paragraph, possibly including supplementary figure 4 into the text; in this regard, authors could invert the two panels (A and B) in order to first provide a useful summary of the different biomarkers emerged in this work (panel B), and subsequently deepen the aspect linked to their biological role.

   Response: Thank you for this suggestion. These two paragraphs have now been combined under a new subheading under the results section titled “Alignment of biomarkers with the Hallmarks of Cancer” (p16, line 390). We have also changed the panel order of the previous Supplementary Figure 4 and brought it into the main text. This is now Figure 1. The figure itself has also been reformatted for enhanced clarity.

6. I would suggest making the Discussion more concise and focused.

   Response: We have removed unnecessary text and trimmed words where possible to make this more concise.

   Overall, this is an interesting review, providing a lot of information which, in my opinion, should be better organized to facilitate understanding of the most relevant biomarkers.

   Response: Thanks for your positive comments. We have reorganised the narrative which we believe has helped improve the flow and assist understanding.

Editorial comments:

1. Please provide: one preferred degree qualification per author and indicate any full professors; affiliation details (department, institute, city, state, country) for each author; full institutional correspondence address for corresponding author.

   Response: All this information has been added.

2. Please check that all author details and affiliations are correct in both the main text and appendix investigator lists (if applicable). We do not guarantee that we will fix errors or omissions after publication (if your article is accepted).

   Response: All checked for accuracy

3. Please add a conflict of interest statement that matches the ICMJE forms, which were submitted with your first draft. Authors should be referred to by their initials in this section. If there are none, then please state "The authors declared no conflicts of interest" or "The other authors declared no conflicts of interest".

   Response: The following has been added: “VS received funding from the University of Aberdeen Development Trust and NHS Grampian Endowments. The other authors declared no conflicts of interest”.

4. Please add a contributors section, detailing specifically what each author did in the preparation of this manuscript. These statements should match those in your author statement forms, which were submitted with your first draft.

   Response: This information has been added

5. We require confirmation that the paper has not been submitted to another journal and has not been published in whole or in part elsewhere previously.

   Response: All authors confirm that this paper has not been submitted to another journal and has not been published in whole or in part elsewhere previously

6. For papers listed in references that are "in press" we need to see a galley proof and letter from the publisher stating that it is 'in press' as well as the full expected citation (ie, publication date/volume/issue etc).

   Response: None of the papers cited are “in press”

7. Images that have been published previously should be accompanied by a statement indicating permission to reproduce the image. If required, further assistance can be obtained from the editorial team. If you have borrowed published images from colleagues, you must obtain permission from the publisher of the paper, not just from the authors. If all the figures are your own and have not been published before then this requirement does not apply.

   Response: All images were generated by the authors

8. The maximum length of a Review is 4500 words.

   Response: Excluding title, abstract, references and Tables, the manuscript is 4497 words

9. The maximum number of references is 75. Please cut your reference list. As a last resort, references can be moved to an appendix, however, the appendix must be cited in the main text at a relevant place and a statement to the effect of “reference for further reading can be found in the appendix” must be added.

   Response: We have limited the references in the main-text to 100 and these are in the reference list. References 101-239 are in the Web Appendix.
   The start of the reference list in the main text states: “We cited 239 references in this manuscript, including the 197 studies that met the inclusion criteria of the systematic review. The first 100 references are listed below with the rest in the Appendix (Page 130).”
   The start of reference list in the Web Appendix states: “*Continued from the main text. References 1-100 are listed in the main text*”.
   We hope this is acceptable and can revise again if required.

10. References should be in the Vancouver style and numbered in the order in which they first appear in the manuscript. If the references "move" from the body text into tables or figures,

please maintain the sequence of citation. Please ensure tables and figures are cited correctly in the body text to prevent the need for renumbering of references should the table and figure citations subsequently move. Please ensure that reference numbering throughout the manuscript is not inserted with electronic referencing software, such as Endnote.

Response: References have been added in the Vancouver style and we have ensured that the numbering is consistent throughout the manuscript. All references are in plain-text format and have not been inserted using a referencing software.

11. If your paper is a systematic review, please check our 'Systematic reviews and meta-analyses formatting guidelines' to ensure that your paper is formatted correctly. Please note that you will need to provide a PRISMA flowchart if so.

    Response: A PRISMA flowchart is provided as Figure S1 (appendix page 124). We have rephrased the narrative to make this clearer.

12. Please supply the webappendix as a single PDF file, with the pages paginated - when you refer to an item in the appendix, please refer to the page number on which it appears, not the table or section. Please note that we will be unable to correct any errors in the webappendix, including errors or omissions in author names or affiliations, following publication; as such, please check carefully when submitting.

    Response: Noted and actioned. The Web Appendix now includes a Table of contents to help find the supplementary information more easily.

13. Please state whether any authors are employed by NIH.

    Response: None of the authors are NIH employees

Other editorial changes made by the authors

To stay within the word count we have abbreviated hormone receptor to HR and reduced some parts of the text. These changes are indicated in red font.

Tense changed from ‘predict’ to ‘predicted’ in section Morphological and/or phenotypic features and ‘shows’ to ‘showed’ in discussion. Replaced ‘molecules’ with ‘biomarkers’ in discussion.

Removed ‘the’ in first section of Genetic and Transcriptomic Markers and changed ‘was’ to ‘were’ (also highlighted in red font).

Removed ‘index’ in the section Other proteomic markers as this word was repeated.

Tables 1-5 in the main text have been reformatted to minimise blank areas in some of the columns (highlighted in red font).