

Field cancerization in breast cancer

Emanuela Gadaleta¹ , Graeme J Thom¹ , Helen Ross-Adams¹ , Louise J Jones^{2*} and Claude Chelala^{1*} 

¹ Centre for Cancer Biomarkers and Biotherapeutics, Barts Cancer Institute, Queen Mary University of London, London, UK

² Centre for Tumour Biology, Barts Cancer Institute, Queen Mary University of London, London, UK

*Correspondence to: C Chelala, Centre for Cancer Biomarkers and Biotherapeutics, Barts Cancer Institute, Queen Mary University of London, Charterhouse Square, London, EC1M 6BQ, UK or LJ Jones, Centre for Tumour Biology, Barts Cancer Institute, Queen Mary University of London, Charterhouse Square, London, EC1M 6BQ, UK. E-mail: c.chelala@qmul.ac.uk or ljjones@qmul.ac.uk

Abstract

Breast cancer affects one in seven women worldwide during their lifetime. Widespread mammographic screening programs and education campaigns allow for early detection of the disease, often during its asymptomatic phase. Current practice in treatment and recurrence monitoring is based primarily on pathological evaluations but can also encompass genomic evaluations, both of which focus on the primary tumor. Although breast cancer is one of the most studied cancers, patients still recur at a rate of up to 15% within the first 10 years post-surgery. Local recurrence was originally attributed to tumor cells contaminating histologically normal (HN) tissues beyond the surgical margin, but advances in technology have allowed for the identification of distinct aberrations that exist in the peri-tumoral tissues themselves. One leading theory to explain this phenomenon is the field cancerization theory. Under this hypothesis, tumors arise from a field of molecularly altered cells that create a permissive environment for malignant evolution, which can occur with or without morphological changes. The traditional histopathology paradigm dictates that molecular alterations are reflected in the tissue phenotype. However, the spectrum of inter-patient variability of normal breast tissue may obfuscate recognition of a cancerized field during routine diagnostics. In this review, we explore the concept of field cancerization focusing on HN peri-tumoral tissues: we present the pathological and molecular features of field cancerization within these tissues and discuss how the use of peri-tumoral tissues can affect research. Our observations suggest that pathological and molecular evaluations could be used synergistically to assess risk and guide the therapeutic management of patients.

© 2022 The Authors. *The Journal of Pathology* published by John Wiley & Sons Ltd on behalf of The Pathological Society of Great Britain and Ireland.

Keywords: breast cancer; field cancerization; cancer-adjacent tissues; histologically normal

Received 14 February 2022; Revised 23 March 2022; Accepted 29 March 2022

No conflicts of interest were declared.

Introduction

Breast cancer is the most common invasive cancer in women worldwide, affecting one in seven women during their lifetime [1]. Widespread mammographic screening programs, enhanced education campaigns, and advances in detection methods and therapeutic regimes allow for the detection of asymptomatic disease, resulting in significant improvements in overall survival [2]. This early detection has led to the increased implementation of breast conservation therapy, which has significantly improved the management of early breast cancer.

Currently, pathological evaluations and genomic tests that focus on the primary tumor are employed to determine therapeutic management and risk of recurrence. Although breast cancer has been studied extensively, patients who undergo breast conservation surgery with post-operative radiotherapy still have a recurrence rate of between 3% and 15% within 10 years [3].

Originally, it was presumed that local recurrence following surgery was attributable to tumor cells contaminating histologically normal (HN) tissues beyond the surgical margin [4]. However, advances in technology have allowed researchers to better characterize tumors and their associated HN tissues, with several studies reporting aberrations in peri-tumoral tissues and determining that these tissues provide distinct information beyond that available from tumor. Two main theories have been proposed to explain these observations: the field cancerization theory and the tumor microenvironment theory [5,6].

The field cancerization theory proposes the creation of a field of molecularly altered cells or an environment with an inherent predisposition to malignant evolution that can occur with or without morphological change [7]; the tumor microenvironment theory suggests that microenvironmental factors are key drivers of tumor initiation and progression. These theories are unlikely to be

mutually exclusive but instead co-exist in an intricate relationship that potentially influences manifested phenotypes [8].

In this review, we explore the concepts of field cancerization in HN tissues. We present the pathologic and molecular characteristics of field cancerization and discuss the implications of using peri-tumoral tissues in research (Figure 1).

Development and evolution of field cancerization theories

‘Field cancerization’ was first defined by Slaughter *et al* to describe hyperplastic and atypical epithelia surrounding multifocal oral squamous carcinoma [7]. They attributed the high rate of local recurrence to a progressive change in peri-tumoral cells rather than the influence of pre-existing cancer cells. The authors concluded that a preconditioned microenvironmental milieu was created, in which cells underwent irreversible changes to confer tumor growth.

Since this original definition, the concept of field cancerization has been continually adapted to remain abreast of knowledge attained from advances in molecular technologies [9–11]. In light of evidence from genetic progression models, current definitions support the presupposition that cells within the preneoplastic field accumulate pro-tumorigenic molecular alterations. However, the tissue itself may not exhibit obvious morphological change as determined by histopathology, meaning that these aberrant cells remain clinically hidden. The criteria that define morphologically normal tissues are being challenged as we begin to understand how ‘normal variation’ may impact cancer risk.

The mechanisms by which the molecularly altered cells develop are not fully understood, but clonal expansion and intra-epithelial migration of these cells within contiguous epithelial structures have been proposed [11]. It has been suggested that cancerization can be attributed to the divergence, selection, and expansion of one or more clones that ultimately confer a survival advantage over the resident cell population [9].

The ‘etiologic field effect’ reassesses the concept of field cancerization by taking a more holistic approach [11]. Here, it is suggested that endogenous and exogenous factors, such as ageing, diet, smoking, and exposure to ultraviolet light, individually and together affect body tissues and contribute to a field of cancer predisposition.

More recently, it has been proposed that field cancerization could be considered in terms of phenotypic traits, rather than underlying specific molecular aberrations. Here, the cancerized field is defined as a collection of cells that have acquired a subset of phenotypic alterations required for malignancy even in the absence of histological alterations [9].

The theories of field cancerization will continue to evolve as studies discern the cellular mechanisms

underpinning its aetiology, thus allowing for the demarcation of cancerized fields and elucidation of the properties that define patient outcome and response to therapy.

Determinants of field cancerization in breast cancer

Technological advances offer the opportunity to gain a deeper understanding of cancer biology and its underlying mechanisms from both a pathological and a molecular aspect. Numerous studies have shown that patches of molecularly altered, cancer-primed, cells are present in HN tissues surrounding breast tumors and that these cells create an environment with a predisposition to progress to cancer [7,12–20]. These studies offer unparalleled insights into the hidden depths of these peri-tumoral tissues.

Pathological

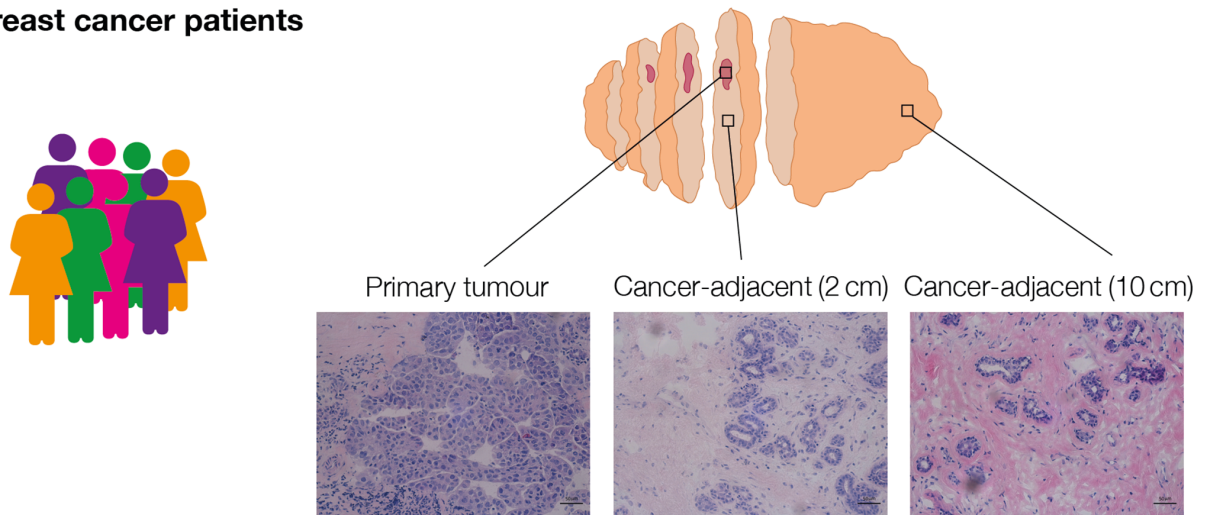
The central tenet of histopathology is that molecular alterations are reflected in tissue phenotype, and decades of research have related specific phenotypic changes to disease behaviour. In the breast, there is a well-defined series of proliferative changes that confer a variable risk for development of breast cancer [21], with the risk being highest for atypical ductal hyperplasia [22].

Molecular studies on atypical ductal hyperplasia have demonstrated some shared genomic alterations with established malignancy [17,23], which supports its precursor nature and representation of cancerized field. Recognition of these entities and subsequent management of patients based on predicted levels of associated risk are the mainstay of protocol-driven patient management. However, beyond these well-recognized benign and atypical lesions, there is almost infinite variation in the histological appearance of ‘normal’ breast tissue as well as more subtle changes not captured in routine diagnostics that may reflect a field cancerization effect. This morphological variation represents the phenotypic manifestation of an individual’s genetic and hormonal environment and is increasingly recognized to influence cancer risk.

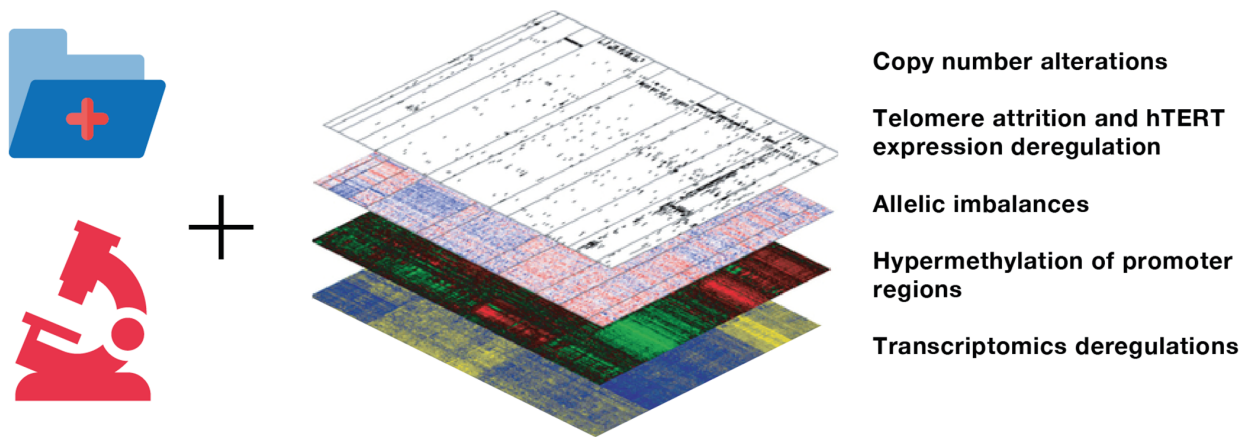
One such histological variation is manifested by the physiological process of lobular involution [24,25], whereby there is a reduction in the number of acini in a terminal ductal lobular unit (TDLU) and a reduction in the number of lobules per unit area. The occurrence of lobular involution is most prominent around the perimenopausal years but is highly variable between individuals, in terms of both stage of development and extent of development.

The status of lobular involution has been assessed semi-quantitatively across large patient cohorts and demonstrated that more advanced lobular involution in women who have had a biopsy showing benign breast disease is associated with significantly reduced breast cancer risk [25,26]. Furthermore, longitudinal assessment of lobular involution in women undergoing multiple biopsies indicates that delayed lobular involution is associated with breast cancer risk. This normal variation

A Breast cancer patients



B Data linkage



C Stratification by clinical and molecular features of cancer-adjacent tissues



Figure 1. Overview of how healthcare records and individual molecular data could be used to guide stratified care. (A) Patients would be referred to hospital, where a subset would be presented with a breast cancer diagnosis. For each patient, the molecular characteristics of the excised tumor and their associated peri-tumoral tissues would be recorded in their medical records. (B) Information from both electronic health records (e.g. demographics, pathology, imaging, etc.) and molecular assessments of tumor and surrounding histologically normal tissues (e.g. genomics, transcriptomics, etc.) would be available to clinicians. (C) This information would then be used to stratify patients into clinically significant groups, such as risk of recurrence, or used to help guide the clinical management of patients based on pharmacogenomic vulnerabilities within the residing tissues.

in breast histology also appears to modify the risk generated by recognized proliferative lesions. For example, in women with atypical epithelial hyperplasia, patients with no lobular involution in surrounding breast tissue

show a significantly higher risk of developing breast cancer compared with those with complete lobular involution, with a relative risk of 7.79 versus 1.49, respectively [27].

Another normal variation in breast tissue composition is reflected in mammographic density. It is well established that high mammographic density, reflecting an increased ratio of fibroglandular elements to adipose tissue, is a major risk factor for breast cancer [28]. Studies suggest increased collagen deposition to be a major contributor to mammographic density [29], with enhanced peri-ductal collagen alignment, resulting in increased tissue stiffening, potentially mediating the pro-tumorigenic effects of mammographic density [30].

While mammographic density represents a global reflection of breast cancer risk, the direct impact of collagen density and organization on tumor initiation and progression had been shown in several model systems [31,32], and roles for stromal activation and epithelial–mesenchymal transition (EMT) have been widely implicated in tumorigenesis [33,34]. As with lobular involution, subtle features of stromal structure are not captured in routine diagnostic pathology yet may significantly influence subsequent cancer risk.

Linking both mammographic density and lobular involution is the effect of estrogen. Mammographic density is influenced by reproductive and menstrual factors, with pregnancy and the menopause associated with a decrease in density [35]. Similarly, exogenous hormone use also influences breast density [36]. Studies correlating endogenous hormone levels and breast density are inconsistent [37,38], probably due to multiple confounding factors, though a relationship is supported by the association between some SNPs in the *ESR1* gene and percent mammographic density [39]. Similarly, in both pre- and post-menopausal women, higher circulating levels of estrogen metabolites are associated with significantly higher TDLU counts and reduced lobular involution [40]. These findings support the well-recognized association between high estrogen levels and breast cancer risk and suggest that at least part of this risk is mediated through modulation of breast phenotype.

The growing digitization of histopathology and advancements in machine learning approaches offer the opportunity to capture and quantitate a greater wealth of information from histological sections, as was shown recently with the application of a convolutional neural network for automated quantitation of lobular involution in breast biopsies [41]. This opens the potential to better assess field cancerization features in apparently normal tissue and improve patient care.

Molecular

The risk of breast cancer recurrence is determined primarily by histopathological evaluations of the tumor, the results of which are used to guide the therapeutic management of patients. However, genetic, epigenetic, and transcriptomic alterations have been identified in cancer-adjacent tissues of many epithelial cancers, including head and neck [7,12], colorectal [13,42–44], skin [45,46], bladder [13,47–49], lung [14,50], prostate [8,13,51,52], ovarian [53,54], and breast [10,13–15,18–20,55–68]. Moreover, persistent exposure to

environmental risk factors (e.g. tobacco carcinogens and diet) and infectious agents (e.g. human papilloma virus and *Helicobacter pylori* infections), and the presence of physiological conditions (e.g. diabetes), can further induce molecular insults in cancer-adjacent tissues, resulting in an environment primed for tumor development [42,69–71].

Evidence suggests that HN tissue resected adjacent to primary tumors can be prognostic [13,72,73]. Furthermore, characterization of the spatial landscape of the cancerized field has shown that molecular aberrations exist in morphologically normal tissues excised up to 24 cm from the primary tumor, and that the field itself is composed of distinct profiles of plasticity [55]. These observations suggest that assessing tumors in combination with their surrounding tissues could improve prognostic and therapeutic determinations.

Pan-cancer studies provide a comprehensive overview of the mechanistic profiles and aberrant pathways shared in peri-tumoral tissues across cancer cohorts, such as enrichment in immune response/inflammation, metabolism, and cell growth [13,14,48]. These large-scale studies use sequencing and SNP data available from The Cancer Genome Atlas (TCGA). While being an invaluable source of large-scale sample data generated from an array of platforms (e.g. RNA-seq, DNA-seq, methylation arrays, CGH, and SNP arrays, etc.) and sample types (e.g. tissues from tumor, metastasis, and cancer-adjacent tissues as well as blood-derived normal), TCGA does not provide information pertaining to the margins of excision, thereby limiting extrapolations of the spatial implications of the cancerized field. Pan-cancer studies are a robust method to highlight alterations shared across cancer types, but they can lack the granularity offered by cancer-specific research.

Genomic instability is a major characteristic of most cancers that has also been observed in peri-tumoral tissues in breast cancer [61,64,66]. Defining events of genomic instability include telomere attrition, allelic imbalance (including loss of heterozygosity), promoter hypermethylation, copy number alterations (CNAs), and somatic mutations.

Table 1 presents an overview of key molecular markers reported to be associated with field cancerization in breast cancer.

Genetic and epigenetic aberrations in breast cancer

Telomere attrition and allelic imbalance

These have been identified in both tumor and matched cancer-adjacent tissues [10,55,59,61,67,74]. Telomeres in paired HN tissue resected within 1 cm from the tumor margin have been reported to be significantly shorter than in paired peri-tumoral tissues resected at 5 cm [10]. Similarly, the frequency of allelic imbalance and loss of heterozygosity is reported to decrease with increased distance from tumor [68]. The cancer-adjacent allelic imbalance profiles revealed substantial

Table 1. Molecular alterations reported to be associated with field cancerization in breast cancer.

Molecular alteration	Summary of observations	Specimens studied	Reference
Genomics: telomere attrition	Telomere attrition was observed in both cancer cells and luminal epithelial cells in HN TDLUs adjacent to cancer	Tumor Matched cancer-adjacent tissue	Kurabayashi <i>et al</i> [74]
Genomics: telomere attrition and unbalanced loci	Telomere attrition and unbalanced loci were observed in HN cancer-adjacent tissues, with this genomic instability being a function of distance from tumor	Matched HN tissue excised 1 cm from tumor margin (TAHN-1) Matched HN tissue excised 5 cm from tumor margin (TAHN-5) Healthy breast tissue from reduction mammoplasty	Heaphy <i>et al</i> [10]
Genomics: allelic imbalance	Allelic imbalance in matched HN microdissected breast TDLUs from breast cancer patients and <i>BRCA1</i> mutation carriers was three-fold greater than in the reduction mammoplasty control group	HN microdissected breast TDLUs from sporadic breast cancer patients (precise distance from tumor is indeterminate) HN microdissected breast TDLUs from <i>BRCA1</i> gene mutation carriers HN microdissected breast TDLUs from reduction mammoplasty	Larson <i>et al</i> [59]
Genomics: allelic imbalance	The mean frequency of allelic imbalance was higher in HN tissue adjacent to the primary cancer (15.4%) relative to distant tissue from the same breast (3.7%)	Tumor (laser-assisted microdissection) DCIS (laser-assisted microdissection) Matched disease-free tissue adjacent to the primary tumor (laser-assisted microdissection) Matched HN distant tissue (laser-assisted microdissection)	Ellsworth <i>et al</i> [68]
Genomics: loss of heterozygosity	Normal breast epithelial cells obtained from women whose risk of breast cancer had been calculated using the Gail model identified associations between frequency of loss of heterozygosity and patient risk, with a lower risk score of 16.7% reported in patients without loss of heterozygosity, compared with 22.9% if loss of heterozygosity is present	Fine needle aspirations from asymptomatic women with known Gail risk score	Euhus <i>et al</i> [63]
Genomics: differential methylation at CpG loci	Significantly more CpG loci were identified as differentially methylated between contralateral-normal and tumor (63 271 CpG loci $q < 0.01$) than between ipsilateral-normal and tumor (38 346 CpG loci $q < 0.01$). Furthermore, differential methylation in ipsilateral-normal relative to contralateral-normal tissue (9562 CpG loci $p < 0.01$) was also observed	Tumor Matched ipsilateral HN tissue excised ≥ 3 cm from tumor margin Matched contralateral HN tissue from contralateral breast	Muse <i>et al</i> [65]
Genomics: genome-wide DNA methylation and copy number calls	Identified hypervariable levels of DNA methylation and copy number alterations (CNAs) in normal-adjacent tissue (relative to tissue from healthy breast samples), which then became further enriched in the matched tumors. Furthermore, changes in DNA methylation in normal cells were more predictive of breast cancer status than their CNV counterparts	Tumor Unmatched cancer-adjacent tissue excised ≥ 3 cm from tumor margin Healthy breast from reduction mammoplasty Tumor (TCGA) Matched cancer-adjacent tissue excised ≥ 2 cm from tumor margin (TCGA)	Gao <i>et al</i> [64]
Genomics: hypermethylation of <i>RUNX3</i>	Normal tissue in close proximity to the primary tumor exhibited hypermethylation of <i>RUNX3</i>	Tumor Matched HN tissue excised 1 cm from visible tumor boundary (N1) Matched HN tissue excised 2 cm from visible tumor boundary (N2) Matched HN tissue excised 3 cm from visible tumor boundary (N3) Matched HN tissue excised 4 cm from visible tumor boundary (N4) Healthy breast tissue from reduction mammoplasty	Cheng <i>et al</i> [75]
Genomics: hypermethylation and allelic imbalance of <i>RASSF1A</i> promoter region	Four locations (1, 2, 3, and 4 cm) in HN breast tissue from the affected and contralateral breast of breast cancer patients identified hypermethylation in the ipsilateral samples relative to the contralateral with the effect being more pronounced in the vicinity of tumor	Tumor Matched HN tissue excised 1 cm from tumor margin (TAHN-1) Matched HN tissue excised 2 cm from tumor margin (TAHN-2) Matched HN tissue excised 3 cm from tumor margin (TAHN-3) Matched HN tissue excised 4 cm from tumor margin (TAHN-4) Healthy breast tissue from reduction mammoplasty	Yan <i>et al</i> [61]

(Continues)

Table 1. Continued

Molecular alteration	Summary of observations	Specimens studied	Reference
Genomics: promoter methylation	<i>RARB2</i> was observed hypermethylated with cancer-adjacent tissues (32%) relative to unaffected breast (9%). Promoter methylation of <i>RASSF1A</i> and <i>APC</i> occurred more frequently in breast tissues from unaffected women at high risk for breast cancer than in tissues from women at low/intermediate risk of breast cancer	Tumor (fine needle aspiration) Matched benign tissue (fine needle aspiration). Samples taken from ipsilateral breast in quadrant opposite to cancer, with precise distance from tumor indeterminate Unaffected patients (fine needle aspiration)	Lewis et al [19]
Genomics: copy number alterations and loss of heterozygosity	Identification of tissue-specific CNAs in cancer-adjacent tissues from solid tumors, including breast invasive carcinoma	Tumor (TCGA) Matched cancer-adjacent tissue excised ≥ 2 cm from tumor margin (TCGA)	Jakubek et al [48]
Genomics: copy number alterations	Study of 282 females with sporadic breast cancer identifies 108 patients (38.3%) with cancer-associated CNAs in at least one aberrant cancer-free breast tissue	Tumor Matched HN tissue (Polish cohort, Bydgoszcz) excised 4–8 cm from tumor margin Matched HN tissue (Swedish cohort, Falun) excised at variable distances from tumor margin Matched HN tissue (Polish cohort, Krakow) excised at variable distances from tumor margin (maximum 15 cm) Matched HN (tissue Polish cohort, Gdansk) excised at variable distances from tumor margin	Forsberg et al [55]
Genomics and transcriptomics: profiling	Breast epithelial samples obtained from ducts leading to breast carcinomas and matched samples from ducts on the opposite side of the nipple. Determined increased mRNA perturbation in proximity to the primary tumor, with these aberrations not being explained by CNAs	Healthy breast tissue from reduction mammoplasty Tumor Matched HN tissue taken from duct between tumor and nipple (D1) Matched epithelial sample closest to tumor with some samples exhibiting atypical hyperplasia taken from duct between tumor and nipple (D2) Matched HN control taken from the duct on the other side of the nipple within the same breast (O1)	Abdalla et al [66]
Genomics and transcriptomics: profiling	About 40% of HN cancer-adjacent tissues harbored genomic defects in DNA copy number (10%), sequence, methylation, or in RNA sequence (>40%). These molecular alterations were not associated with significant differences in overall survival	Tumor (TCGA) Matched cancer-adjacent tissue excised ≥ 2 cm from tumor margin (TCGA)	Troester et al [72]
Transcriptomics and proteomics: profiling	Transcriptomic and proteomic analysis of breast tumors and matched HN tissues resected proximal to (<2 cm) and distant from (5–10 cm) the primary tumor, using tissues from reduction mammoplasties as baseline. Four distinct transcriptomic subtypes are identified within matched normal tissues: immune; matrisome/EMT, non-coding enriched and metabolic, with the latter associated with poor prognosis ($p < 0.001$, HR 6.1)	Tumor Matched HN tissue resected <2 cm from primary tumor (proximal) Matched HN tissue resected 5–10 cm from primary tumor (distal) Prophylactic mastectomy Healthy breast tissue from reduction mammoplasty Tumor (TCGA) Matched cancer-adjacent tissue excised ≥ 2 cm from tumor margin (TCGA)	Gadaleta et al [57]
Transcriptomics: profiling	Pan-cancer study in which profiles from cancer-adjacent tissues were deemed to represent an intermediate state between tumor and healthy, with both tumor-associated and unique features observed in cancer-adjacent tissues.	Tumor (TCGA) Matched cancer-adjacent excised ≥ 2 cm from tumor margin (TCGA) Healthy breast from autopsy (GTEx)	Aran et al [13]
Transcriptomics: profiling	Pan-cancer study in which profiles from cancer-adjacent tissues were found to provide more information about patient survival than tumors	Tumor (TCGA) Matched HN excised ≥ 2 cm from tumor margin (TCGA)	Huang et al [14]
Transcriptomics: molecular subtypes	Transcriptomic analysis of two breast cancer datasets observed intrinsic tumor subtypes to be reflected in HN cancer-adjacent tissues, with this observation not dependent on distance from primary tumor	Tumor (PWBCS cohort) Matched HN tissue excised <2 cm from tumor margin (PWBCS cohort) Tumor (TCGA) Matched cancer-adjacent tissue excised ≥ 2 cm from tumor margin (TCGA) Matched HN tissue excised <2 cm from tumor margin (peri-tumoral, NBS cohort) Matched HN tissue excised >2 cm from tumor margin (remote, NBS cohort)	Casbas-Hernandez et al [62]
Transcriptomics: profiling		Tumor	

(Continues)

Table 1. Continued

Molecular alteration	Summary of observations	Specimens studied	Reference
	Transcriptomic analyses stratified profiles from HN tissues adjacent to breast cancer or ductal carcinoma <i>in situ</i> (DCIS) data into two extratumoral subtypes: <i>active</i> and <i>inactive</i> , with the <i>active</i> subtype significantly associated with overall survival in estrogen receptor (ER)-positive patients (HR 2.5, $p = 0.062$) and hormone-treated patients (HR 2.6, $p = 0.045$)	DCIS Matched HN tissue excised <2 cm from tumor/DCIS margin (peri-tumoral) Matched HN tissue excised >2 cm from tumor/DCIS margin (remote)	Román-Pérez <i>et al</i> [60]
Transcriptomics: <i>TERT</i> gene expression	Normal tissue proximal to breast tumors were found to contain a population of human mammary epithelial cells (HMECs) that expressed human telomerase reverse transcriptase (hTERT) expression levels similar to HMECs within the tumor. hTERT expression decreased in HMECs from HN tissues at increased distance from tumor	Tumor Matched HN tissue excised 1 cm from tumor margin (TAHN-1) Matched HN tissue excised 3 cm from tumor margin (TAHN-3) Matched HN tissue excised 5 cm from tumor margin (TAHN-5)	Trujillo <i>et al</i> [15]
Transcriptomics: profiling	Comparison of the gene expression profiles of microdissected HN epithelium from tissues of patients with breast cancer, at high risk of breast cancer, and from reduction mammoplasty identified defining features within the HN tissues of patients with breast cancer relative to reduction mammoplasty	Healthy breast tissue from reduction mammoplasty HN epithelium from breast cancer patients HN epithelium from prophylactic mastectomy HN epithelium from reduction mammoplasty	Graham <i>et al</i> [58]
Transcriptomics: profiling	Transcriptomic profiling and immunohistochemistry discerned a signature of differential gene expression that discriminated between paired breast tissues excised at resection margins of 1 and 5 cm from the primary tumor	Tumor Matched HN tissue excised 1 cm from tumor margin (TAHN-1) Matched HN tissue excised 5 cm from tumor margin (TAHN-5)	Trujillo <i>et al</i> [73]
Transcriptomics: profiling	Gene expression profiling of normal-appearing TDLUs of ER-positive breast cancer patients ($n = 14$) and of reduction mammoplasty ($n = 15$) identified 105 differentially expressed genes. Investigations identified cancer-associated alterations in the normal-appearing TDLUs relative to TDLUs from reduction mammoplasty	Healthy breast tissue from reduction mammoplasty HN microdissected HN breast TDLUs adjacent to untreated ER-positive breast cancer HN microdissected HN breast TDLUs from reduction mammoplasty	Tripathi <i>et al</i> [76]

concordance with those of tumor, implying clonal evolution [10,56].

The Gail model is a predictive breast cancer risk assessment algorithm that uses demographic and clinical data to estimate a woman's absolute risk of developing breast cancer within specific time frames (up to the age of 90 years), with a score exceeding 1.67% being defined as high risk [77–79]. Euhus *et al* conducted a study in which 30 asymptomatic women, whose risk of breast cancer had been calculated using the Gail model (11 with normal cytology and 19 with proliferative cytology), underwent breast epithelial cell sampling via fine-needle aspirate. Associations between the frequency of loss of heterozygosity and cancer risk were reported, with the patient's lifetime risk score increasing from 16.7% to 22.9% if loss of heterozygosity was present [63].

Promoter hypermethylation

Aberrant DNA methylation has been observed in tissues from tumor and matched HN tissues, with the methylomic landscape of cancer-adjacent tissue being

epigenetically distinct from that of tissue extracted from the contralateral breast [65]. This observation is supported by inter-individual studies in which the cancer-adjacent tissues are found to be more highly methylated compared with healthy control reduction mammoplasty samples [64].

Promoter region hypermethylation has been observed in both cancer-adjacent tissues and *in situ* breast carcinomas [18,19,61,64,75]. Hypermethylation of the *RASSF1A* promoter was observed in both tumor and cancer-adjacent tissues. Separately, hypermethylation of promoters, including *RASSF1A*, was found to be indicative of increased risk of breast cancer: present in 70% of benign breast tissues from women at high risk (Gail risk index ≥ 2) compared with 29% of women at low/intermediate risk [63].

Copy number aberrations (CNAs)

While CNAs have been identified in HN tissues adjacent to cancer, the prevalence and degree of these perturbations in peri-tumoral tissues are not as extensive as other genomic and transcriptomic events and they do not appear to confer predictive capabilities

[55,61,64,66,72]. Furthermore, cancer-adjacent tissues primarily exhibit tumor-associated CNA profiles, with gains in established oncogenes and growth factor receptors, such as *ERBB2*, *EGFR*, and *FGFR1* [48,55].

Mutations

Cancer-associated somatic mutation profiles appear more frequent relative to CNAs in cancer-adjacent tissues. Troester *et al* reported 25% of cancer-adjacent samples in TCGA to have moderate-to-high levels of tumor-like somatic mutations relative to CNAs in 10% in the triplet samples analysed, although the variant allele fraction was low (typically less than 5%), which was consistent with matched tumor cellularity [72]. Similarly, mutation hotspots in specific genes, including *ZNF143*, *ALDOA*, and *LEPROTL1*, have also been shown to reflect proximity to the primary tumor [66].

The occurrences of genomic instability presented suggest that local peri-tumoral tissues provide a permissive environment for tumorigenesis and progression, with attenuation of this effect with increased distance from the primary tumor.

Transcriptomic aberrations in breast cancer

Cancer-adjacent tissues are often considered to represent an intermediate state between tumor and healthy breast tissue, with their transcriptomic profiles sharing features of both. However, distinct gene expression characteristics defining HN tissues and their associated subtypes suggest that they represent a distinct entity [13,15,57,58,60,62,72].

Global transcriptomic analyses have linked genes differentially expressed in HN peri-tumoral tissues to a range of deregulated cancer-associated pathways, including wound healing, extracellular matrix remodelling, altered metabolism, and EMT [13,15,57,58,60,72]. Specific prognostic signatures have consequently been developed [5,57,60,72]. Román-Pérez *et al* identified distinct *active* and *inactive* transcriptomic subtypes in cancer-adjacent tissues from estrogen receptor-positive patients, with the former correlating with poor outcome [60]. This *active* mRNA/miRNA signature was also identified in 40% of matched cancer-adjacent breast tissues ($n = 142$) from the TCGA breast cohort, with this subtype associated with worse poor 10-year survival in estrogen receptor-positive patients [72]. Moreover, analysis of TCGA data revealed that gene expression profiles of peri-tumoral tissues correlate more closely with clinical outcome measures than their corresponding tumors [14].

Several transcriptomic studies also report that the number and degree of aberrations are dependent on distance from the primary tumor [13,15,66], in concordance with findings on genomic instability. However, we identified four distinct transcriptomic subtypes within matched peri-tumoral tissues excised adjacent to (<2 cm) and distal from (5–10 cm) the primary breast tumor, which were found to be independent of distance from tumor. Most importantly, we found the *metabolic* subtype, characterized by deregulation of mediators involved in metabolic processes, lipid and cholesterol

metabolism, and hypoxia-related events, to be significantly associated with poor prognosis [hazard ratio (HR) 6.1]. We also observed a *matrisome/EMT* subtype that was enriched in matrisomal elements; however, the small sample size and differing sampling design of the TCGA validation cohort prevented meaningful conclusions as to the clinical implications of this group.

That stromal alterations may contribute to a field cancerization effect is supported by Trujillo *et al*, who analysed HN breast tissue taken 1 and 5 cm from invasive breast cancer and compared the gene expression profile with reduction mammoplasty tissue [73]. They identified a gene expression signature reflecting extracellular matrix remodelling, fibrosis, and EMT, which was more prevalent in tissues closest to the tumor and absent in reduction mammoplasty tissue. Similarly, the *active* miRNA/RNA subtype identified in HN tissues represented processes associated with activated stroma or EMT [60,72].

Taken together, these studies suggest that the molecular profiles of cancer-adjacent tissues both reflect those of the tumor and possess distinct intrinsic features. Significant heterogeneity in the cancerized field is observed, with clear clinically relevant molecular events linked directly to prognosis occurring even distal from the primary tumor.

Further study is required to determine the extent to which molecular aberrations contribute to the preconditioned milieu for tumorigenesis. With genomic and transcriptomic events being inherently intertwined, a prognostic tool incorporating a combination of these events will likely provide the greatest predictive value.

Challenges for research

Access to well-annotated specimens from biobanks

Well-annotated tumor and peri-tumoral samples with comprehensive clinical data are key to fully understanding the landscape of field cancerization. Here, the role of biobanks is crucial. Historically, these were localized specimen repositories set up for the requirements of specific research projects. As the infrastructures developed, biobanks soon became national and international resources designed to accelerate translational research efforts. Specimens supplied by biobanks are the cornerstone to generating molecular data and in allowing biobanks to evolve away from a narrow specimen-focused approach towards a data-driven future. However, for biobanks to reach their full potential, they must provide access to high-quality clinical samples linked to comprehensive clinical, spatial, and molecular information.

One biobank that has recognized and anticipated the importance of complete annotation and data legacy is the Breast Cancer Now Tissue Bank (BCNTB) [80]. The BCNTB ensures that pathologists and researchers work side-by-side to provide well-annotated specimens from an extensive range of tissues and body fluids. These include tumors and tissues taken at specified distances

from margins of resection (adjacent <2 cm and surrounding 5–10 cm), which allow researchers to map the spatial characteristics of the cancerized field. Furthermore, the BCNTB has implemented a data return policy in which researchers are required to return research data generated from projects using BCNTB samples, meaning that each set of specimens gains additional layers of molecular and pathological data. This then becomes available to subsequent researchers, allowing the knowledge to build up incrementally, and is available to query from a dedicated BCNTB Analytics Hub [81].

Biobanks need to provide an active response to the changing needs of researchers and anticipate specimen and data requirements, thus ensuring that these are exploited to their full clinical potential.

Coding of field cancerization concepts in electronic health records (EHRs)

The ability to access and interpret all information within EHRs is a prerequisite to understand disease processes, and to tailor treatments and health services effectively and safely for patient benefit. NHS England has announced that it will fund an initiative to audit metastatic breast cancer to provide accurate information about patients living with this disease and reform how metastatic breast cancer is currently coded in EHRs [82], which will increase the understanding of recurrence and the patient's clinical journey. While EHRs currently include information pertaining to the status of the resection margins, there are few molecular data available relating to this margin and its surrounding tissues.

Genomics initiatives, such as the 100,000 Genomes Project, allow for a new branch of genomic medicine to serve the NHS, helping to realize the promise of personalized medicine and driving cutting-edge research. If field cancerization is responsible for recurrence, then limiting sequencing to tumors alone could prove a missed opportunity to elucidate whether the cancerized field arises from tumor influence, or whether it is, itself, an independent entity that drives tumorigenesis and recurrence. Introducing coding concepts, both histological and molecular, for field cancerization into EHRs to complement histological assessments could drive new opportunities for patient-centered research and information.

Balancing the benefits and harm of defining a high-risk cancerized field

Breast cancer screening programs enable the detection of asymptomatic breast cancer, allowing for early intervention and resulting in reductions of up to 20% in absolute mortality in the women screened [83]. However, screening programs exacerbate overdiagnosis, detecting indolent cancers that would not have given rise to clinically relevant disease [83–85].

Estimates of overdiagnosis vary from near zero to 50%, with these patients subject to unnecessary treatment, increased psychological stress, and potential adverse reactions and complications associated with

treatment. Variations in these estimates are dependent on the definition of overdiagnosis applied. While the numerator represents the number of patients overdiagnosed, the denominator varies between studies depending how this value is defined: by screening alone; during the whole screening period; during both the screening and the interval period; or during the screening period and for the remainder of the individual's lifetime [83,84].

Longitudinal studies, integrating molecular, pathological, and epidemiological data with corresponding patient clinical records, offer researchers the opportunity to characterize molecularly aberrant fields that generate a pro-tumorigenic environment. They could also help identify patients, in whom these fields are present, who will progress to cancer, which will be essential to minimize overdiagnosis. Furthermore, these investigations also have the potential to provide insights into whether the defining molecular and phenotypic characteristics of a cancerized field categorized as high risk can be modulated by modifications to environmental factors.

The added value of molecular pathological epidemiology

To better understand the cancerized field and its implication for breast cancer, researchers need to explore beyond the boundaries of traditional methods. Novel lines of evidence point towards the benefits of integrating molecular pathology and epidemiology – termed molecular pathological epidemiology – to enable researchers to link factors such as environmental exposures and host genetics to pathologic features [86–91].

Endogenous and exogenous factors (exposures) can contribute to specific disease processes and heterogeneity of neoplastic disease: for instance, antibiotic-induced microbial dysbiosis in the gut is associated with tumor progression [92–96]. Emerging evidence of the relationship between dysbiosis and breast cancer has uncovered microbial signatures providing type-specific communities of organisms distinct to each breast cancer type [97,98]. Differences in breast microbiota DNA profiles from tumors relative to paired cancer-adjacent tissues and healthy controls have also been reported, with the microbial profile of cancer-adjacent tissues exhibiting greater similarities with their corresponding tumor than with healthy controls [93,97,98].

While it is possible that these studies indicate that dysbiosis influences the environment surrounding a tumor, potentially modulating the risk of breast cancer, the reverse may also be true, in that the cancerized field itself creates a niche favoring differential microbiome composition, which in turn may influence progression of the disease.

The field of molecular pathological epidemiology allows for the examination of complex relationships between a compendium of exposures. For instance, positive associations have been reported between increased body mass index, hip-to-waist ratio and reduced physical activity, and increased risk of breast cancer

[99–101]. Similarly, epigenetic modifications have been found to be influenced by environmental exposures in women at high risk of breast cancer relative to healthy controls [12,90,102].

While genetic and environmental exposures are known to contribute to breast cancer risk, the magnitude of interaction between these remains undefined for many genes [103–107]. However, there is evidence that host and environmental factors can modify the penetrance in germline *BRCA1/2* mutation carriers. Increased fat mass and dysmetabolism are recognized to promote breast cancer risk in the general population but there is evidence to indicate a greater impact in *BRCA2* carriers [108], and diet has also been shown to influence *BRCA* penetrance [109]. It has also been shown that smoking for at least 5 years prior to a first pregnancy increases the risk of developing breast cancer in *BRCA1/2* mutation carriers [110]. A recent review outlines the complexity of gene–environment interactions in *BRCA1/2* mutation carriers [111].

Molecular pathological epidemiology studies that integrate data from the clinical records of patients will help to clarify the causal relationships between environmental exposures, pathologic interpretations, molecular aberrations, and disease evolution within the cancerized field.

Biobanks can facilitate such research by implementing protocols to gather information pertaining to lifestyle and environmental exposures, as well as procedures for the collection, management, storage, and distribution of longitudinal microbiota specimens.

Designating appropriate control cohorts

Collecting a germline genetic profile from buffy coat samples is common practice in DNA sequencing studies. However, it is less common to have matched tissue control samples for comparative transcriptomics, where spatial and temporal effects are significant confounders.

That the field of HN peri-tumoral tissues may have undergone a molecular insult is an inherent problem in study design. Using cancer-adjacent tissues to represent a ‘healthy’ control could provide a false representation of baseline gene expression, leading to inaccurate tumor profiles, as well as limiting the opportunity to order tumor-associated molecular alterations chronologically.

Designation of an appropriate control cohort is dependent on the objective of a study. While cancer-adjacent samples may have acquired molecular alterations, use of contralateral tissues could prove invaluable when conducting intra-individual and site-specific comparisons. Furthermore, pan-cancer studies have reported that changes in gene expression between paired tumor and adjacent tissues improve the forecast of disease aetiology and patient outcome more than tumor alone [14].

The use of a control cohort of breast tissue from women without breast cancer (nor at known increased risk of breast cancer), such as tissue from reduction mammoplasties, is often preferred. However, their use introduces inherent variability into analyses, particularly when attempting to determine subtle patient-specific

variations. This heterogeneity can be minimized by adopting stringent pre-defined inclusion criteria, such as demographic (age, ethnicity, gender) or clinical (age of menarche, menopausal status, parity, BMI) and lifestyle features (smoking, contraceptive use), matching those of the breast cancer cohort.

It is important to elucidate alterations within the cancerized field as well as their clinical implications because it is these cancer-adjacent tissues that will remain unresected in commonly accepted guidelines and in breast conservation therapy. Researchers need to assess the strengths and weaknesses of each candidate set of controls, to make an informed decision and delineate appropriate baseline cohorts. In addition, they also need to have access to specimens or data with complete clinical and surgical information from independent cohorts for validation.

Exploiting the peri-tumoral field for patient benefit

Natural processes associated with ageing, such as methylation and a lifetime exposure to endogenous and exogenous factors, cause genomic changes that resemble those of early cancer [9,71,112,113]. However, treating the whole body as a cancerized field has limited clinical utility. Similarly, germline breast cancer susceptibility gene variants, such as *BRCA1* and *BRCA2*, are not considered to contribute towards the traditional definition of a cancerized field [11].

While the cancerized field represents a promising therapeutic target, it is important to acknowledge that presence of the field alone will not determine progression to cancer or propensity to metastasize. To avoid the pitfalls of overdiagnosis and overtreatment, we need to understand the underlying mechanism and molecular events and be able to differentiate between changes that drive predisposed lineages further down the tumorigenic evolutionary pathway and those changes that have minimal cancerization effect.

Currently, invasive procedures, such as colonoscopy and cervical cytology, tend to be employed for the early detection of cancer. Non-invasive detection strategies usually rely on detecting cancer-specific biomarkers, which are developed for their ability to differentiate between tumor and normal tissues. Many of these strategies detect the presence of late-stage cancer or metastasis, with few able to detect early-stage disease [114–116].

Two major liquid biopsy studies, CancerSEEK and GRAIL, support the premise that circulating tumor DNA is a promising biomarker for diagnostic screening for sporadic cancer. CancerSEEK analysed mutated DNA and eight standard protein biomarkers in 1005 patients previously diagnosed with stage I–III colorectal, breast, gastric, liver, oesophageal, ovarian, and pancreatic cancers [114]. GRAIL applied targeted whole-genome bisulphite sequencing of plasma DNA to identify distinct methylation patterns associated with specific cancers to detect a number of those cancers early and

simultaneously provide information about the organ of origin [116]. This prospective multi-center case-control study comprised 6689 individuals, split one third with cancer and two thirds without cancer, representing more than 50 primary cancer types across all clinical stages. While both studies exhibit specificities of more than 99%, sensitivity was proven to be dependent on stage (CancerSEEK, ~40% in stage I disease to approximately 80% in stage III disease; GRAIL, ~18% in stage I disease to ~91% in stage IV disease).

Circulating tumor DNA has been shown to recapitulate a subset of mutational signatures identified in the primary tumor and to have good correlation with tumor burden in solid tumors [117–122]. Breast cancer studies have reported elevated circulating tumor DNA burden to be a strong predictive marker of disease progression, decreased progression-free and overall survival, and poor response to treatment [118,123–126]. However, almost no studies into solid tumors have attempted to detect distinct cancer-adjacent profiles in liquid biopsies.

Wu *et al* conducted a study on 27 patients with head and neck squamous carcinoma [12]. The authors observed mutations unique to cancer-adjacent tissues in post-operative liquid biopsies from blood and saliva samples from a subset of patients. Moreover, the study reported that integrating data on both cancer-adjacent-specific mutations and tumor-specific mutations increased the sensitivity of post-operative monitoring to predict relapse. These observations raise the question of whether treatment and predictive models could benefit from insights gained from distinct signatures originating from the cancerized field.

Studies indicating the importance of HN cancer-adjacent tissues in the context of their ability to prognosticate outcome and response to treatment highlight the requirement for further investigations [12,57,63]. Understanding the intricacies of field cancerization and developing predictive biomarkers that exploit the characteristics of this cancerized field could help to improve prognostic and therapeutic determinations and inform clinical decision-making.

It is important to appreciate that field cancerization is dynamic and influenced by a range of genetic and environmental factors over time. As such, longitudinal clinical observation of breast cancer patients would enable clinicians to monitor trends in blood biomarker profiles and determine when the patient's risk of recurrence changes. Clinical management of the patient is thus enhanced by introducing more information to guide the level of surveillance.

Diagnostic tissue biopsies serve to provide a snapshot of the tumor and surrounding tissues at a given point in time. Repeated invasive biopsies are not a feasible solution to monitoring patient disease trajectory, whereas liquid biopsies have potential as clinical modalities able to yield important diagnostic, prognostic, and therapeutic information and to reduce the dependency on invasive procedures and radiological tests.

The demarcation of surgical margins is not consistent between patients. Molecular signatures developed to

detect high-risk peri-tumoral tissues could be used in conjunction with morphological examinations to define excision margins and determine the completeness of excision in breast conservation therapy. Evidence of clear pathological and molecular margins could indicate decreased risk of recurrence, thereby sparing patients additional radiation therapy. Similarly, evidence of a field biomarker profile associated with increased risk of cancer recurrence could distinguish patients requiring increased surveillance or more aggressive treatment at the time of surgical intervention. As such, it is important for cancerized field-related concepts to be recorded as a part of medical health records.

Conclusion

Continued technological advances will undoubtedly have significant repercussions in elucidating the cancerized field. This may result in a paradigmatic shift from histopathological-driven therapeutics to a holistic evaluation of clinical and molecular characteristics to complement current diagnostic techniques and help inform clinical decision-making.

Acknowledgements

We would like to thank all the patients who donated their tissues, without whom advances in research would not be possible. EG is funded by Breast Cancer Now (BCNTB Award to LJJ and CC). GJT and HR-A are funded by Barts Cancer Institute (Incentivisation Award to CC).

Author contributions statement

EG and CC conceptualized the manuscript. EG, GJT and LJJ wrote the manuscript. EG, GJT, CC, HR-A and LJ reviewed and edited the manuscript. EG, HR-A, GJT and CC made revisions.

References

1. Cancer Research UK (CRUK). Breast cancer statistics. [Accessed 14 February 2022]. Available from: <https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/breast-cancer>.
2. Duffy SW, Tabár L, Yen AMF, *et al*. Mammography screening reduces rates of advanced and fatal breast cancers: results in 549,091 women. *Cancer* 2020; **126**: 2971–2979.
3. Cleveland Clinic. Breast cancer recurrence: symptoms & treatment. [Accessed 14 February 2022]. Available from: <https://my.clevelandclinic.org/health/diseases/8328-breast-cancer-recurrence>.
4. Sadanandam A, Lal A, Benz SC, *et al*. Genomic aberrations in normal tissue adjacent to *HER2*-amplified breast cancers: field cancerization or contaminating tumor cells? *Breast Cancer Res Treat* 2012; **136**: 693–703.

5. Lebya K, Garcia-Smith R, Swaminathan R, et al. Towards a personalized surgical margin for breast conserving surgery – implications of field cancerization in local recurrence. *J Surg Oncol* 2017; **115**: 109–115.
6. Troester MA, Lee MH, Carter M, et al. Activation of host wound responses in breast cancer microenvironment. *Clin Cancer Res* 2009; **15**: 7020–7028.
7. Slaughter DP, Southwick HW, Smejkal W. Field cancerization in oral stratified squamous epithelium; clinical implications of multicentric origin. *Cancer* 1953; **6**: 963–968.
8. Trujillo KA, Jones AC, Griffith JK, et al. Markers of field cancerization: proposed clinical applications in prostate biopsies. *Prostate Cancer* 2012; **2012**: 302894.
9. Curtius K, Wright NA, Graham TA. An evolutionary perspective on field cancerization. *Nat Rev Cancer* 2018; **18**: 19–32.
10. Heaphy CM, Bisoffi M, Fordyce CA, et al. Telomere DNA content and allelic imbalance demonstrate field cancerization in histologically normal tissue adjacent to breast tumors. *Int J Cancer* 2006; **119**: 108–116.
11. Lochhead P, Chan AT, Nishihara R, et al. Etiologic field effect: reappraisal of the field effect concept in cancer predisposition and progression. *Mod Pathol* 2015; **28**: 14–29.
12. Wu P, Xie C, Yang L, et al. The genomic architectures of tumour-adjacent tissues, plasma and saliva reveal evolutionary underpinnings of relapse in head and neck squamous cell carcinoma. *Br J Cancer* 2021; **125**: 854–864.
13. Aran D, Camarda R, Odegaard J, et al. Comprehensive analysis of normal adjacent to tumor transcriptomes. *Nat Commun* 2017; **8**: 1077.
14. Huang X, Stern DF, Zhao H. Transcriptional profiles from paired normal samples offer complementary information on cancer patient survival – evidence from TCGA pan-cancer data. *Sci Rep* 2016; **6**: 20567.
15. Trujillo KA, Hines WC, Vargas KM, et al. Breast field cancerization: isolation and comparison of telomerase-expressing cells in tumor and tumor adjacent, histologically normal breast tissue. *Mol Cancer Res* 2011; **9**: 1209–1221.
16. Heaphy CM, Griffith JK, Bisoffi M. Mammary field cancerization: molecular evidence and clinical importance. *Breast Cancer Res Treat* 2009; **118**: 229–239.
17. Gao Y, Niu Y, Wang X, et al. Genetic changes at specific stages of breast cancer progression detected by comparative genomic hybridization. *J Mol Med (Berl)* 2009; **87**: 145–152.
18. Lehmann U, Länger F, Feist H, et al. Quantitative assessment of promoter hypermethylation during breast cancer development. *Am J Pathol* 2002; **160**: 605–612.
19. Lewis CM, Cler LR, Bu D-W, et al. Promoter hypermethylation in benign breast epithelium in relation to predicted breast cancer risk. *Clin Cancer Res* 2005; **11**: 166–172.
20. Rheinbay E, Parasuraman P, Grimsby J, et al. Recurrent and functional regulatory mutations in breast cancer. *Nature* 2017; **547**: 55–60.
21. Elston CW, Ellis IO (eds). Epithelial hyperplasia. In *Systemic Pathology, The Breast*, Symmers WStC (Emeritus Editor), Volume 13. Churchill Livingstone: Edinburgh, 1998; 65–89.
22. Kader T, Hill P, Rakha EA, et al. Atypical ductal hyperplasia: update on diagnosis, management, and molecular landscape. *Breast Cancer Res* 2018; **20**: 39.
23. O’Connell P, Pekkel V, Fuqua SA, et al. Analysis of loss of heterozygosity in 399 premalignant breast lesions at 15 genetic loci. *J Natl Cancer Inst* 1998; **90**: 697–703.
24. Henson DE, Tarone RE. Involution and the etiology of breast cancer. *Cancer* 1994; **74**: 424–429.
25. Milanese TR, Hartmann LC, Sellers TA, et al. Age-related lobular involution and risk of breast cancer. *J Natl Cancer Inst* 2006; **98**: 1600–1607.
26. Baer HJ, Collins LC, Connolly JL, et al. Lobule type and subsequent breast cancer risk: results from the Nurses’ Health Studies. *Cancer* 2009; **115**: 1404–1411.
27. Radisky DC, Visscher DW, Frank RD, et al. Natural history of age-related lobular involution and impact on breast cancer risk. *Breast Cancer Res Treat* 2016; **155**: 423–430.
28. Martin LJ, Boyd NF. Mammographic density. Potential mechanisms of breast cancer risk associated with mammographic density: hypotheses based on epidemiological evidence. *Breast Cancer Res* 2008; **10**: 201.
29. Li T, Sun L, Miller N, et al. The association of measured breast tissue characteristics with mammographic density and other risk factors for breast cancer. *Cancer Epidemiol Biomarkers Prev* 2005; **14**: 343–349.
30. McConnell JC, O’Connell OV, Brennan K, et al. Increased periductal collagen micro-organization may contribute to raised mammographic density. *Breast Cancer Res* 2016; **18**: 5.
31. Levental KR, Yu H, Kass L, et al. Matrix crosslinking forces tumor progression by enhancing integrin signaling. *Cell* 2009; **139**: 891–906.
32. Provenzano PP, Eliceiri KW, Campbell JM, et al. Collagen reorganization at the tumor–stromal interface facilitates local invasion. *BMC Med* 2006; **4**: 38.
33. Schäfer M, Werner S. Cancer as an overhealing wound: an old hypothesis revisited. *Nat Rev Mol Cell Biol* 2008; **9**: 628–638.
34. Kaushik S, Pickup MW, Weaver VM. From transformation to metastasis: deconstructing the extracellular matrix in breast cancer. *Cancer Metastasis Rev* 2016; **35**: 655–667.
35. Vachon CM, Kuni CC, Anderson K, et al. Association of mammographically defined percent breast density with epidemiologic risk factors for breast cancer (United States). *Cancer Causes Control* 2000; **11**: 653–662.
36. Greendale GA, Reboussin BA, Slone S, et al. Postmenopausal hormone therapy and change in mammographic density. *J Natl Cancer Inst* 2003; **95**: 30–37.
37. Boyd N, Martin L, Chavez S, et al. Breast-tissue composition and other risk factors for breast cancer in young women: a cross-sectional study. *Lancet Oncol* 2009; **10**: 569–580.
38. Jung S, Stanczyk FZ, Eggleston BL, et al. Endogenous sex hormones and breast density in young women. *Cancer Epidemiol Biomarkers Prev* 2015; **24**: 369–378.
39. Fjeldheim FN, Frydenberg H, Flote VG, et al. Polymorphisms in the estrogen receptor alpha gene (*ESR1*), daily cycling estrogen and mammographic density phenotypes. *BMC Cancer* 2016; **16**: 776.
40. Oh H, Khodr ZG, Sherman ME, et al. Relation of serum estrogen metabolites with terminal duct lobular unit involution among women undergoing diagnostic image-guided breast biopsy. *Horm Cancer* 2016; **7**: 305–315.
41. de Bel T, Litjens G, Ogony J, et al. Automated quantification of levels of breast terminal duct lobular (TDLU) involution using deep learning. *NPJ Breast Cancer* 2022; **8**: 13.
42. Del Puerto-Nevado L, Minguez P, Corton M, et al. Molecular evidence of field cancerization initiated by diabetes in colon cancer patients. *Mol Oncol* 2019; **13**: 857–872.
43. Guo H, Zeng W, Feng L, et al. Integrated transcriptomic analysis of distance-related field cancerization in rectal cancer patients. *Oncotarget* 2017; **8**: 61107–61117.
44. Hawthorn L, Lan L, Mojica W. Evidence for field effect cancerization in colorectal cancer. *Genomics* 2014; **103**: 211–221.
45. Cornejo CM, Jambusaria-Pahlajani A, Willenbrink TJ, et al. Field cancerization: treatment. *J Am Acad Dermatol* 2020; **83**: 719–730.
46. Hu B, Castillo E, Harewood L, et al. Multifocal epithelial tumors and field cancerization from loss of mesenchymal CSL signaling. *Cell* 2012; **149**: 1207–1220.
47. Czerniak B, Dinney C, McConkey D. Origins of bladder cancer. *Annu Rev Pathol* 2016; **11**: 149–174.

48. Jakubek YA, Chang K, Sivakumar S, *et al.* Large-scale analysis of acquired chromosomal alterations in non-tumor samples from patients with cancer. *Nat Biotechnol* 2020; **38**: 90–96.
49. Thomsen MBH, Nordentoft I, Lamy P, *et al.* Comprehensive multi-regional analysis of molecular heterogeneity in bladder cancer. *Sci Rep* 2017; **7**: 11702.
50. Kadara H, Fujimoto J, Yoo SY, *et al.* Transcriptomic architecture of the adjacent airway field cancerization in non-small cell lung cancer. *J Natl Cancer Inst* 2014; **106**: dju004.
51. Amirrad F, Pytak PA, Sadeghiani-Pelar N, *et al.* Prostate field cancerization and exosomes: association between CD9, early growth response 1 and fatty acid synthase. *Int J Oncol* 2020; **56**: 957–968.
52. Mehrotra J, Varde S, Wang H, *et al.* Quantitative, spatial resolution of the epigenetic field effect in prostate cancer. *Prostate* 2008; **68**: 152–160.
53. Buller RE, Skilling JS, Sood AK, *et al.* Field cancerization: why late “recurrent” ovarian cancer is not recurrent. *Am J Obstet Gynecol* 1998; **178**: 641–649.
54. Giannopoulou L, Chebouti I, Pavlakis K, *et al.* RASSF1A promoter methylation in high-grade serous ovarian cancer: a direct comparison study in primary tumors, adjacent morphologically tumor cell-free tissues and paired circulating tumor DNA. *Oncotarget* 2017; **8**: 21429–21443.
55. Forsberg LA, Rasi C, Pekar G, *et al.* Signatures of post-zygotic structural genetic aberrations in the cells of histologically normal breast tissue that can predispose to sporadic breast cancer. *Genome Res* 2015; **25**: 1521–1535.
56. Försti A, Louhelainen J, Söderberg M, *et al.* Loss of heterozygosity in tumour-adjacent normal tissue of breast and bladder cancer. *Eur J Cancer* 2001; **37**: 1372–1380.
57. Gadaleta E, Fourgoux P, Pirrò S, *et al.* Characterization of four subtypes in morphologically normal tissue excised proximal and distal to breast cancer. *NPJ Breast Cancer* 2020; **6**: 38.
58. Graham K, de las Morenas A, Tripathi A, *et al.* Gene expression in histologically normal epithelium from breast cancer patients and from cancer-free prophylactic mastectomy patients shares a similar profile. *Br J Cancer* 2010; **102**: 1284–1293.
59. Larson PS, de las Morenas A, Cerda SR, *et al.* Quantitative analysis of allele imbalance supports atypical ductal hyperplasia lesions as direct breast cancer precursors. *J Pathol* 2006; **209**: 307–316.
60. Román-Pérez E, Casbas-Hernández P, Pirone JR, *et al.* Gene expression in extratumoral microenvironment predicts clinical outcome in breast cancer patients. *Breast Cancer Res* 2012; **14**: R51.
61. Yan PS, Venkataramu C, Ibrahim A, *et al.* Mapping geographic zones of cancer risk with epigenetic biomarkers in normal breast tissue. *Clin Cancer Res* 2006; **12**: 6626–6636.
62. Casbas-Hernandez P, Sun X, Roman-Perez E, *et al.* Tumor intrinsic subtype is reflected in cancer-adjacent tissue. *Cancer Epidemiol Biomarkers Prev* 2015; **24**: 406–414.
63. Euhus DM, Cler L, Shivapurkar N, *et al.* Loss of heterozygosity in benign breast epithelium in relation to breast cancer risk. *J Natl Cancer Inst* 2002; **94**: 858–860.
64. Gao Y, Widschwendter M, Teschendorff AE. DNA methylation patterns in normal tissue correlate more strongly with breast cancer status than copy-number variants. *EBioMedicine* 2018; **31**: 243–252.
65. Muse ME, Titus AJ, Salas LA, *et al.* Enrichment of CpG island shore region hypermethylation in epigenetic breast field cancerization. *Epigenetics* 2020; **15**: 1093–1106.
66. Abdalla M, Tran-Thanh D, Moreno J, *et al.* Mapping genomic and transcriptomic alterations spatially in epithelial cells adjacent to human breast carcinoma. *Nat Commun* 2017; **8**: 1245.
67. Deng G, Lu Y, Zlotnikov G, *et al.* Loss of heterozygosity in normal tissue adjacent to breast carcinomas. *Science* 1996; **274**: 2057–2059.
68. Ellsworth DL, Ellsworth RE, Love B, *et al.* Genomic patterns of allelic imbalance in disease free tissue adjacent to primary breast carcinomas. *Breast Cancer Res Treat* 2004; **88**: 131–139.
69. Braakhuis BJM, Tabor MP, Kummer JA, *et al.* A genetic explanation of Slaughter’s concept of field cancerization: evidence and clinical implications. *Cancer Res* 2003; **63**: 1727–1730.
70. Hattori N, Ushijima T. Epigenetic impact of infection on carcinogenesis: mechanisms and applications. *Genome Med* 2016; **8**: 10.
71. Takeshima H, Ushijima T. Accumulation of genetic and epigenetic alterations in normal cells and cancer risk. *NPJ Precis Oncol* 2019; **3**: 7.
72. Troester MA, Hoadley KA, D’Arcy M, *et al.* DNA defects, epigenetics, and gene expression in cancer-adjacent breast: a study from The Cancer Genome Atlas. *NPJ Breast Cancer* 2016; **2**: 16007.
73. Trujillo KA, Heaphy CM, Mai M, *et al.* Markers of fibrosis and epithelial to mesenchymal transition demonstrate field cancerization in histologically normal tissue adjacent to breast tumors. *Int J Cancer* 2011; **129**: 1310–1321.
74. Kurabayashi R, Takubo K, Aida J, *et al.* Luminal and cancer cells in the breast show more rapid telomere shortening than myoepithelial cells and fibroblasts. *Hum Pathol* 2008; **39**: 1647–1655.
75. Cheng ASL, Culhane AC, Chan MWY, *et al.* Epithelial progeny of estrogen-exposed breast progenitor cells display a cancer-like methylome. *Cancer Res* 2008; **68**: 1786–1796.
76. Tripathi A, King C, de la Morenas A, *et al.* Gene expression abnormalities in histologically normal breast epithelium of breast cancer patients. *Int J Cancer* 2008; **122**: 1557–1566.
77. Gail MH, Brinton LA, Byar DP, *et al.* Projecting individualized probabilities of developing breast cancer for white females who are being examined annually. *J Natl Cancer Inst* 1989; **81**: 1879–1886.
78. Gail MH, Costantino JP, Pee D, *et al.* Projecting individualized absolute invasive breast cancer risk in African American women. *J Natl Cancer Inst* 2007; **99**: 1782–1792.
79. Matsuno RK, Costantino JP, Ziegler RG, *et al.* Projecting individualized absolute invasive breast cancer risk in Asian and Pacific Islander American women. *J Natl Cancer Inst* 2011; **103**: 951–961.
80. Breast Cancer Now Tissue Bank (BCNTB). [Accessed 16 April 2022]. Available from: <https://breastcancernow.org/breast-cancer-research/breast-cancer-now-tissue-bank>.
81. Gadaleta E, Pirrò S, Dayem Ullah AZ, *et al.* BCNTB bioinformatics: the next evolutionary step in the bioinformatics of breast cancer tissue banking. *Nucleic Acids Res* 2018; **46**: D1055–D1061.
82. Breast Cancer Now Media. [Accessed 16 April 2022]. Available from: <https://breastcancernow.org/about-us/media/press-releases/breast-cancer-now-achieves-landmark-moment-nhs-england-announces-national-metastatic-breast-cancer-audit>.
83. Independent UK Panel on Breast Cancer Screening. The benefits and harms of breast cancer screening: an independent review. *Lancet* 2012; **380**: 1778–1786.
84. Bulliard J-L, Beau A-B, Njor S, *et al.* Breast cancer screening and overdiagnosis. *Int J Cancer*. 2021; **149**: 846–853.
85. Molassiotis A, Tyrovolas S, Giné-Vázquez I, *et al.* Organized breast cancer screening not only reduces mortality from breast cancer but also significantly decreases disability-adjusted life years: analysis of the Global Burden of Disease Study and screening programme availability in 130 countries. *ESMO Open* 2021; **6**: 100111.
86. Herrington CS, Poulsom R, Coates PJ. Recent Advances in Pathology: The 2019 Annual Review Issue of *The Journal of Pathology*. *J Pathol* 2019; **247**: 535–538.
87. Hamada T, Nowak JA, Milner DA Jr, *et al.* Integration of microbiology, molecular pathology, and epidemiology: a new paradigm to explore the pathogenesis of microbiome-driven neoplasms. *J Pathol* 2019; **247**: 615–628.
88. Ogino S, Nishihara R, VanderWeele TJ, *et al.* Review Article: The role of molecular pathological epidemiology in the study of

- neoplastic and non-neoplastic diseases in the era of precision medicine. *Epidemiology* 2016; **27**: 602–611.
89. Ogino S, Lochhead P, Chan AT, et al. Molecular pathological epidemiology of epigenetics: emerging integrative science to analyze environment, host, and disease. *Mod Pathol* 2013; **26**: 465–484.
 90. Hughes LAE, Simons CCJM, van den Brandt PA, et al. Lifestyle, diet, and colorectal cancer risk according to (epi)genetic instability: current evidence and future directions of molecular pathological epidemiology. *Curr Colorectal Cancer Rep* 2017; **13**: 455–469.
 91. Mima K, Kosumi K, Baba Y, et al. The microbiome, genetics, and gastrointestinal neoplasms: the evolving field of molecular pathological epidemiology to analyze the tumor–immune–microbiome interaction. *Hum Genet* 2021; **140**: 725–746.
 92. Ingman WV. The gut microbiome: a new player in breast cancer metastasis. *Cancer Res* 2019; **79**: 3539–3541.
 93. Urbaniak C, Gloor GB, Brackstone M, et al. The microbiota of breast tissue and its association with breast cancer. *Appl Environ Microbiol* 2016; **82**: 5039–5048.
 94. Xuan C, Shamonki JM, Chung A, et al. Microbial dysbiosis is associated with human breast cancer. *PLoS One* 2014; **9**: e83744.
 95. McKee AM, Kirkup BM, Madgwick M, et al. Antibiotic-induced disturbances of the gut microbiota result in accelerated breast tumor growth. *iScience* 2021; **24**: 103012.
 96. Kirkup MB, McKee A, Makin AK, et al. Perturbation of the gut microbiota by antibiotics results in accelerated breast tumour growth and metabolic dysregulation. *bioRxiv* 2019; doi.org/10.1101/553602 [Not peer reviewed].
 97. Banerjee S, Tian T, Wei Z, et al. Distinct microbial signatures associated with different breast cancer types. *Front Microbiol* 2018; **9**: 951.
 98. Zhang J, Xia Y, Sun J. Breast and gut microbiome in health and cancer. *Genes Dis* 2020; **8**: 581–589.
 99. Papadimitriou N, Dimou N, Tsilidis KK, et al. Physical activity and risks of breast and colorectal cancer: a Mendelian randomisation analysis. *Nat Commun* 2020; **11**: 597.
 100. Barone I, Giordano C, Bonofiglio D, et al. The weight of obesity in breast cancer progression and metastasis: clinical and molecular perspectives. *Semin Cancer Biol* 2020; **60**: 274–284.
 101. Gao C, Patel CJ, Michailidou K, et al. Mendelian randomization study of adiposity-related traits and risk of breast, ovarian, prostate, lung and colorectal cancer. *Int J Epidemiol* 2016; **45**: 896–908.
 102. Delgado-Cruzata L, Wu HC, Liao Y, et al. Differences in DNA methylation by extent of breast cancer family history in unaffected women. *Epigenetics* 2014; **9**: 243–248.
 103. Rudolph A, Milne RL, Truong T, et al. Investigation of gene–environment interactions between 47 newly identified breast cancer susceptibility loci and environmental risk factors. *Int J Cancer* 2015; **136**: E685–E696.
 104. Nickels S, Truong T, Hein R, et al. Evidence of gene–environment interactions between common breast cancer susceptibility loci and established environmental risk factors. *PLoS Genet* 2013; **9**: e1003284.
 105. Barrdahl M, Canzian F, Joshi AD, et al. Post-GWAS gene–environment interplay in breast cancer: results from the Breast and Prostate Cancer Cohort Consortium and a meta-analysis on 79,000 women. *Hum Mol Genet* 2014; **23**: 5260–5270.
 106. Schoeps A, Rudolph A, Seibold P, et al. Identification of new genetic susceptibility loci for breast cancer through consideration of gene–environment interactions. *Genet Epidemiol* 2014; **38**: 84–93.
 107. Aschard H, Zaitlen N, Lindström S, et al. Variation in predictive ability of common genetic variants by established strata: the example of breast cancer and age. *Epidemiology* 2015; **26**: 51–58.
 108. Bruno E, Oliverio A, Paradiso A, et al. Lifestyle characteristics in women carriers of BRCA mutations: results from an Italian trial cohort. *Clin Breast Cancer* 2021; **21**: e168–e176.
 109. Pasanisi P, Bruno E. Breast cancer in BRCA mutations carriers: is it time for a “lifestyle” primary prevention? *Epidemiol Prev* 2018; **42**: 369–371.
 110. Li H, Terry MB, Antoniou AC, et al. Alcohol consumption, cigarette smoking, and risk of breast cancer for BRCA1 and BRCA2 mutation carriers: results from The BRCA1 and BRCA2 Cohort Consortium. *Cancer Epidemiol Biomarkers Prev* 2020; **29**: 368–378.
 111. Daniele A, Divella R, Pilato B, et al. Can harmful lifestyle, obesity and weight changes increase the risk of breast cancer in BRCA1 and BRCA2 mutation carriers? A mini review. *Hered Cancer Clin Pract* 2021; **19**: 45.
 112. Li R, Di L, Li J, et al. A body map of somatic mutagenesis in morphologically normal human tissues. *Nature* 2021; **597**: 398–403.
 113. Dotto GP. Multifocal epithelial tumors and field cancerization: stroma as a primary determinant. *J Clin Invest* 2014; **124**: 1446–1453.
 114. Cohen JD, Li L, Wang Y, et al. Detection and localization of surgically resectable cancers with a multi-analyte blood test. *Science* 2018; **359**: 926–930.
 115. Debernardi S, O’Brien H, Alghamdi AS, et al. A combination of urinary biomarker panel and PancRISK score for earlier detection of pancreatic cancer: a case–control study. *PLoS Med* 2020; **17**: e1003489.
 116. Liu MC, Oxnard GR, Klein EA, et al. Sensitive and specific multi-cancer detection and localization using methylation signatures in cell-free DNA. *Ann Oncol* 2020; **31**: 745–759.
 117. Dawson SJ, Tsui DW, Murtaza M, et al. Analysis of circulating tumor DNA to monitor metastatic breast cancer. *N Engl J Med* 2013; **368**: 1199–1209.
 118. Kruger DT, Jansen MPH, Konings IRHM, et al. High ctDNA molecule numbers relate with poor outcome in advanced ER+, HER2–postmenopausal breast cancer patients treated with everolimus and exemestane. *Mol Oncol* 2020; **14**: 490–503.
 119. Lv J, Chen Y, Zhou G, et al. Liquid biopsy tracking during sequential chemo-radiotherapy identifies distinct prognostic phenotypes in nasopharyngeal carcinoma. *Nat Commun* 2019; **10**: 3941.
 120. Moss J, Zick A, Grinshpun A, et al. Circulating breast-derived DNA allows universal detection and monitoring of localized breast cancer. *Ann Oncol* 2020; **31**: 395–403.
 121. Phallen J, Leal A, Woodward BD, et al. Early noninvasive detection of response to targeted therapy in non-small cell lung cancer. *Cancer Res* 2019; **79**: 1204–1213.
 122. Roesch-Ely M, Nees M, Karsai S, et al. Proteomic analysis reveals successive aberrations in protein expression from healthy mucosa to invasive head and neck cancer. *Oncogene* 2007; **26**: 54–64.
 123. Kingston B, Cutts RJ, Bye H, et al. Genomic profile of advanced breast cancer in circulating tumour DNA. *Nat Commun* 2021; **12**: 2423.
 124. Muendlein A, Geiger K, Gaenger S, et al. Significant impact of circulating tumour DNA mutations on survival in metastatic breast cancer patients. *Sci Rep* 2021; **11**: 6761.
 125. Sharma P, Abramson VG, O’Dea A, et al. Clinical and biomarker results from phase I/II study of PI3K inhibitor alpelisib plus nab-paclitaxel in HER2-negative metastatic breast cancer. *Clin Cancer Res* 2021; **27**: 3896–3904.
 126. Tan G, Chu C, Gui X, et al. The prognostic value of circulating cell-free DNA in breast cancer: a meta-analysis. *Medicine (Baltimore)* 2018; **97**: e0197.