Risk factors and novel biomarkers in breast cancer

Thesis submitted for the degree of Doctor of Philosophy

at University College London

by

Evangelia - Ourania Fourkala

Department of Gynaecological Oncology

Institute for Women's Health

University College London

When you set out on your journey to Ithaca, pray that the road is long, full of adventure, full of knowledge. The Lestrygonians and the Cyclops, the angry Poseidon - do not fear them: You will never find such as these on your path, if your thoughts remain lofty, if a fine emotion touches your spirit and your body. The Lestrygonians and the Cyclops, the fierce Poseidon you will never encounter, if you do not carry them within your soul, if your soul does not set them up before you.

Pray that the road is long. That the summer mornings are many, when, with such pleasure, with such joy you will enter ports seen for the first time; stop at Phoenician markets, and purchase fine merchandise, mother-of-pearl and coral, amber, and ebony, and sensual perfumes of all kinds, as many sensual perfumes as you can; visit many Egyptian cities, to learn and learn from scholars. Always keep Ithaca on your mind. To arrive there is your ultimate goal. But do not hurry the voyage at all. It is better to let it last for many years; and to anchor at the island when you are old, rich with all you have gained on the way, not expecting that Ithaca will offer you riches.

Ithaca has given you the beautiful voyage. Without her you would have never set out on the road. She has nothing more to give you.

And if you find her poor, Ithaca has not deceived you. Wise as you have become, with so much experience, you must already have understood what these Ithacas mean.

by Constantine P. Cavafy

(1863 - 1933)

To search gives you purpose,

To find sets you free…

Dedicated to all the women affected by breast cancer

ABSTRACT

Efforts continue to identify and validate novel risk factors / biomarkers for breast cancer and improve current risk prediction models in the general population due to ongoing issues with sensitivity and specificity.

The overall goal of this PhD study is to add to this effort. Specific aims are to (1) examine which is the best source of getting notified for breast cancer diagnosis in the general population since accurate data is crucial for risk assessment studies (2) investigate the association of sex steroids, gonadotrophins and novel assays of sex steroid hormone receptor serum bioactivity (SB) in breast cancer (3) examine whether they can be combined to improve breast cancer risk assessment and (4) identify new DNA methylation markers that might add to such a strategy in the future. To achieve this, a nested case-control study was undertaken within UK Collaborative Trial of Ovarian Cancer Screening.

2629 trial participants were identified via cancer registry (CR) or self-reporting to have breast cancer. Diagnosis was confirmed by the treating clinician. The largest study was undertaken in England and Wales to examine completeness of breast cancer diagnosis within UKCTOCS. Analysis of complete data obtained in 1083 of these women showed CR to be more accurate than self-reporting but associated with time-delays.

Serum samples from 200 eligible breast cancer cases identified through the process and 400 matched-controls were analysed for oestradiol, free-oestradiol, oestrone, androstenedione, testosterone, free-testosterone, progesterone, dehydroepiandrosterone sulphate (DHEAS), sex steroid hormone binding globulin

(SHBG), luteinising hormone (LH), follicle stimulating hormone (FSH) and oestrogen receptor-α and -β and androgen receptor SB. Results showed that sex steroid receptor SB assays could add to breast cancer risk prediction. Additionally, the best oestrogen for breast cancer risk prediction is oestrone and the best androgen is testosterone. High testosterone and FSH levels up to 5 years prior to diagnosis predict breast cancer with high power and may have a synergistic effect.

In a separate case control study of 189 paraffin-embedded breast tissue samples, 55 genes were investigated using MethyLight. DNA methylation alterations were found to be homogeneous in breast cancer with 13 genes being predictive of the disease, suggesting that such changes could be useful as future biomarkers.

Further studies (already underway) involve using high-throughput technology to analyse serum DNA methylation changes and correlate these with the observed serum hormonal changes and build better breast cancer risk prediction models.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

DECLARATION

I declare that this thesis has been composed and the work described in it performed by the candidate Evangelia-Ourania Fourkala. It has not been submitted for another degree either at this or at another university. All sources of information have been acknowledged.

ACKNOWLEDGEMENTS

I would like to express my gratitude and my deepest appreciation to all these people who have supported me to complete this study. In particular I would like to thank:

Professor Usha Menon and Professor Martin Widshwendter my supervisors and mentors; for their support, inspiring leadership and constructive criticism throughout the years that this study took to be completed. For always taking time to answer my questions despite their loads of clinical work and other responsibilities. Their valuable scientific comments, their generosity to share their knowledge and our friendly conversations made this work more pleasant and feasible. For also giving me the opportunity to remain within the team and continue work in such stimulating environment.

Professor Alexey Zaikin and Dr Matthew Burnell for our long conversations, our many meetings and for all these e-mails to help and advice on the statistics required for this thesis. Their help is much appreciated!

All the members of the clinical trial team in the department and in particular Professor Ian Jacobs for giving me the opportunity along with my supervisors to be part of the team, Dr Alexandra Gentry-Maharaj for her guidance with the identification and collection of data on breast cancer cases and Dr Andy Ryan for his help with the follow-up questionnaires. Despite their heavy schedule they were always happy to help. The members in the Tumour Marker Lab, Jeremy Ford and Richard Gunu and the Clinical Biochemistry Department, Dr Ann Dawney and Dr Christina Soromani for sharing their expertise for the hormone measurements and

teaching me all the technical details. My colleagues in DNA Methylation Group, Dr Sophia Apostolidou and Alison Jones for training me with methylation analysis, our coffee break conversations made the work more pleasant.

Our collaborators in Austria and in particular Dr Conny Hauser-Kronberger for providing the paraffin embedded tissue samples and Dr Heidi Fiegl for her contribution with the expression analysis. Our collaborators in Bonn; Professor Hella Lichtenberg-Fraté for accepting me in her team to perform the sex steroid bioactivity assays, apart from being a great mentor during my stay for becoming a great friend. Dr Guido Hasenbrink for teaching me the technique and for being such a good company during the long days in the lab.

All my colleagues, past and present, for creating such a lovely and friendly environment to work; Dr Penny Allen, Nyala Balogun, Gwen Fletcher, Saz Chamberlain, Jenney Sheals, Susan Moffat and Katherine Bailey. A huge list of names could be added with all the amazing people that I met during this project at the Institute for Women's Health.

Many thanks to all the physicians who agreed to complete a questionnaire regarding the histology and treatment of the volunteer women and the nurses who provided histopathology reports.

This work would not been feasible without the funding from Medical Research Council (MRC), and the funds obtained for the UK Collaborative trial of Ovarian Cancer Screening (UKCTOCS) which is mainly supported by MRC, Cancer Research UK, Department of Health, Special Trustees of St Bartholomew's

Hospital, Special Trustees of University College London and the Eve Appeal. Finally, by the Jubilaeumsfonds der Oesterreichischen Nationalbank and a grant from the European Union, Biognosis.

Most importantly, all my thanks to the women throughout the UK who participated in the trial and consented to use their samples for the purposes of the study.

To all my friends in Greece - especially Kerasia and Akis, and to my family and friends in London - especially auntie Soula, Natasa, Anny, Gregoris, Manos, for spicing up my life and for all their support during my academic years.

To Apostoli for his support in the good and bad days. Making my everyday life easier, helping in any aspect so that I could go through with the workload, especially during the period of the write-up when the full time job was taking over my time. He has contributed in so many different levels in my life, making me a happier and better person.

My sister, Margarita along with her family and my brothers, Manolis and Argyris, for always believing in me and for their tremendous support throughout my life. My grandparents, especially Giagia Ourania, and all the other members of the family for always making me feel like a star every time I was visiting my home in Greece.

Last but not least, my parents, Petros and Matoula, for being a role model to me, for all their hard work in order for me to get the best education and upbringing and for always being ‘‘here’’ even though we are thousands of miles apart.

ABBREVIATIONS

AF	Activation function
ANC	Axillary node clearance
AR	Androgen receptor
AUC	Area under the curve
BCQ	Breast cancer questionnaire
BD	Binding Domain
BM	Bisulphite modification
BMI	Body mass index
bp	Base pair
CGIs	CpG islands
CI	Confidence interval
COBRA	Combined bisulphite restriction analysis
CR	Cancer registry
DCIS	Ductal carcinoma *in situ*
DHEAS	Dehydroepiandrosterone sulphate
DHT	Dihydrotestosterone
DNMTs	DNA methyltransferases
EED	Embryonic ectoderm development
EGFR	Epidermal growth factor receptor
ELISA	Enzyme-linked immunosorbent assay
EPIC	European prospective investigation into cancer and nutrition
ER	Oestrogen receptor
ES	Embryonic stem
E_1	Oestrone

E_2	Oestradiol
E_3	Oestriol
EZH2	Enhancer of zeste homologue 2
fE2	Free oestradiol
FL	Fluorescence
FN	False negative
FP	False positive
FSH	Follicle stimulating hormone
fT	Free testosterone
FUQ	Follow-up questionnaire
GFP	Green fluorescence protein
GPs	General practises
HATs	Histone acetylases
hCG	Human chorionic gonadotrophin
HDACs	Histone deacetylases
HDMs	Histone demethylases
HER2	Human epidermal receptor 2
HMTs	Histone methyltransferases
HRE	Hormone response element
HRT	Hormone replacement therapy
ICD	International classification of diseases
IDC	Invasive ductal carcinoma
IQR	Inter-quartile range
LBD	Ligand binding domain
LCIS	Lobular carcinoma *in situ*
LH	Luteinising hormone

MALDI-TOF MS	Matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry
MBDs	Methyl-CpG-binding protein
MRI	Magnetic resonance imaging
MS	Mass spectrometry
MSP	Methylation specific PCR
NHS	National Health Service
NHSCR	NHS cancer registry
OC	Oral Contraceptives
OD	Optical density
ONS	Office of national statistics
OR	Odds ratio
PCGT	Polycomb group target
PCR	Polymerase chain reaction
PMR	Percentage of methylated reference
PPV	Positive predictive value
PR	Progesterone receptor
PRC	Polycomb repressor complex
ROC	Receiver operator characteristics
RT	Reverse Transcription
SAM	S-adenosyl-L-methionine
SB	Serum bioactivity
SELDI-TOF MS	Surface-enhanced laser desorption/ionisation time-of-flight mass spectrometry
SHBG	Sex hormone-binding globulin
SUZ12	Suppressor of zeste homologue 12

TARGIT	Targeted Intraoperative Radiotherapy
TBP	TATA binding protein
TN	True negative
TNM	Staging System: T: tumour size; N: nodes involvement; M: metastasis
TP	True positive
UK	United Kingdom
UKCTOCS	UK collaborative trial of ovarian cancer screening
WLE	Wide local excision
YNB	Yeast nitrogen base

1 INTRODUCTION

1.1 Overall Purpose of the Study

Breast cancer is one of the major health problems facing the world today, being a significant contributor to overall morbidity and mortality. In 2008, the Breast Cancer Campaign Gap Analysis Meeting proposed that in order to make the greatest impact on breast cancer patients the following aspects need to be taken into account: 1) the identification of women predisposed to breast cancer by risk prediction markers and 2) the application of a preventive or early detection strategy [1]. This need is further magnified by the current controversies of the efficacy of breast cancer screening and the concern about over diagnosis and unnecessary treatment [2, 3].

In the last decades, a huge effort has been made to identify risk factors and biomarkers associated with breast cancer that can be used for risk stratification. The most well known to date are age, family history, previous history of benign breast conditions, genetic predisposition, epigenetic changes, reproductive factors and hormonal changes, breast density and environmental influences. Some of these factors have been included in risk prediction models in order to identify women at high risk. At present, the only model that is in clinical use to assess a woman's risk to develop breast cancer is the Gail model [4-9]. However, the chance of correctly classifying a randomly chosen woman with this model is around 59%, only marginally better than chance and whilst statistically significant, the model lacks in sensitivity and specificity. Therefore, more studies are urgently needed to identify and validate risk factors / biomarkers that can effectively stratify women to available prevention / screening strategies [10].

1.2 Project overview

The overall goal of this project is to try to improve breast cancer risk prediction in the general population by adding in the effort for the identification and validation of breast cancer biomarkers. Specific aims were: **1)** a) to identify breast cancer cases in a large cohort – the UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS) which involves 202,638 postmenopausal women aged 50-74 years at recruitment in 2001-5 and b) to examine the accuracy of breast cancer diagnosis within the cohort **2)** a) to examine the association of sex steroid and gonadotrophin hormonal changes in combination with novel assays of sex steroid hormonal receptor bioactivity and b) to investigate whether they can be combined to improve breast cancer risk assessment by examining their synergistic effect in breast carcinogenesis **3)** a) to identify new epigenetic markers for breast cancer and b) to examine the homogeneity of these changes in the disease that might add to such a strategy in the future.

The first objective was to identify breast cancer cases within UKCTOCS using two sources; cancer registry data and self-reporting (UKCTOCS follow-up questionnaire). Confirmation of breast cancer diagnosis and clinicopathological information was collected from the treating physicians in a form of questionnaire specifically designed for the purposes of the study. Based on the fact that collection of accurate cancer diagnosis information has major implications for research studies, especially those that include cancer risk prediction, sensitivity and positive predictive value (PPV) of these two sources individually and combined by comparing it to medical records obtained from the treating physicians were investigated. By the end of this procedure eligible cases for aim 3 were identified based on the clinicopathological data collected from their

treating physicians. Matched controls were selected from within the UKCTOCS cohort. Serum samples donated by eligible participants before breast cancer diagnosis were identified for the cases and with the serum from the identified matched controls were retrieved from the trial bio-bank.

As the association of sex steroid hormones with breast cancer is known for long time, six sex steroid hormones (oestradiol, oestrone, androstenedione, testosterone, dehydroepiandrosterone sulfate (DHEAS) and progesterone) were assayed along with sex hormone-binding globulin (SHBG) and another two sex steroids were calculated (free oestradiol and free testosterone) in order to validate their association with breast cancer risk. Earlier studies had only explored levels of endogenous hormones. Recently sensitive bioactivity assays for steroid hormones which are able to detect minimal levels of hormonal activity have been described. These were used to measure levels of serum bioactivity of oestrogen receptor (ER) -alpha and -beta (ER-α and –β) and androgen receptor (AR) and examine whether they are associated with breast cancer and predict the disease. To better understand the role of gonadotrophins in breast cancer risk, luteinising hormone (LH) and follicle stimulating hormone (FSH) were investigated. The effect of each individual hormone and the bioactivity of their receptors on breast cancer risk were further examined in relation to time of diagnosis (more and less than 2 years before diagnosis). Furthermore, as all previous studies had looked only at the ability of individual hormones to predict breast cancer risk, the markers were combined to investigate if breast cancer risk prediction could be improved in comparison to each individual marker’s effect and to identify any possible synergistic effect.

An attractive alternative for biomarker discovery are epigenetic modifications since they are known to occur early in carcinogenesis. In addition to the above case control study, a separate study of paraffin embedded breast tissue samples from a collaboration with the University of Salzburg, Austria was undertaken to establish whether methylation status could discriminate between non-neoplastic and breast cancer tissue irrespective of whether the DNA has been collected from the centre or the periphery of the breast tumour also addressing the issue of intra-tumour heterogeneity. The epigenetic field defect in breast cancer in non-neoplastic tissue adjacent to breast tumour was also investigated to assess whether DNA methylation status is able to indicate the presence of breast cancer.

1.3 Thesis Design

In chapter 1 which is the introduction of this thesis the overall purpose of the study is described along with the general aim and the thesis design. Chapter 2 includes the literature review. Chapter 3 covers the identification of breast cancer cases within UKCTOCS and completeness of the breast cancer diagnosis from three different sources. Chapter 4 contains the hormonal study looking into the association of sex steroid hormone levels, serum bioactivity of sex steroid hormone receptors and gonadotrophins with breast cancer risk. Chapter 5 covers the studies performed to investigate DNA methylation changes in breast cancer and their possible role as markers for risk prediction. Chapters 3, 4 and 5 contain a short introduction along with materials and methods, results and discussion. Chapter 6 includes summary of the results, future work and conclusions.

2 LITERATURE REVIEW

2.1 Breast Cancer

2.1.1 Epidemiology

In many countries breast cancer is the most common malignancy among women. In UK, there were 39,681 women diagnosed with breast cancer in 2008, accounting for 31% of all female cancers and just over 10,000 women died from the disease in 2000, accounting for 16% of female deaths. A woman's lifetime risk to develop breast cancer is 11%. Majority of the women being diagnosed with breast cancer are postmenopausal. Overall, incidence rates are higher in the developed countries while rates in less developed countries are low but increasing mainly due to changes in life style [11, 12].

2.1.2 Pathology

Breast tumours are almost exclusively adenocarcinomas. Sarcomas and lymphomas are rare and generally excluded from studies of breast cancer. Morphology of the breast tumour is classified by its appearance under the microscope. The site of origin for most of the pathology in the breast is the ducts and the lobules. Historically, most invasive lobular carcinomas were thought to arise within the small terminal ducts of the lobules and ductal carcinomas from the larger or intralobular ducts. However, the distinction between lobular and ductal carcinoma is based more on the histological appearance than on the site of origin and now it is suggested that both types derive from the terminal duct lobuloalveolar unit [13]. Invasive ductal carcinoma (IDC) is the most commonly histological type diagnosed accounting for up to 80% of all breast cancers, and invasive lobular carcinoma (ILC) accounts for 5-15% [14]. It has been suggested

that ILCs are more common in older women and their metastatic pattern is different but they have similar prognosis [15]. There are also some distinct morphologic subtypes such as tubular and medullary, mucinous, papillary and adenoid cystic carcinomas and Paget. These histological types of breast cancer are not that common and have a better prognosis compared to ductal breast cancer [16].

In addition to the invasive cancers, *in situ* carcinomas of the breast are also described including ductal and lobular carcinoma *in situ* (DCIS and LCIS). These non-invasive carcinomas used to be rather uncommon but with the introduction of mammography their diagnosis has increased [17]. By definition, non-invasive breast cancers do not grow through the basement membrane. *In situ* carcinomas can be early precursors of invasive breast cancer. Models have been used to describe the progression from normal healthy breast tissue to atypical ductal hyperplasia to *in situ* carcinoma to breast cancer with studies also suggesting that LCIS can give rise to IDC [18].

2.1.3 Prognosis

Despite the huge number of discovered prognostic and predictive factors for breast cancer, the American Society of Clinical Oncology has concluded that most of them are not satisfactory yet to be recommended for general clinical use. The factors that are currently used include: staging (size of the tumour), lymph node involvement and distant metastasis, grading, ER, progesterone receptor (PR) and human epidermal receptor-2 neu (HER2/neu) expression [19, 20].

Staging refers to the grouping of patients according to the extent of their disease. The first clinical staging system was the Columbia Clinical Classification, which was developed in the 1940-50's. In 2002, a new system was introduced adopted by the International Union against Cancer and the American Joint Committee on Cancer. This system is the one currently in use and is based on the principle of the tumour size (T), nodes involvement (N) and metastases (M) known as TNM [21]. These factors do not predict response to therapy but are only prognostic. Tumour size is defined as the largest diameter of the tumour and is a prognostic factor for breast cancer death regardless of other tumour characteristics. Lymph node involvement is another important independent prognostic factor. Women with lymph node involvement and increasing number of affected lymph nodes are associated with poorer prognosis [22]. Finally, metastasis is a third factor of clinical importance. In general, women with distant metastasis have a median overall survival time of only 2 years [23].

Tumour grade is the classification of the differentiation of the tumour into three groups: low, high and moderate. High grade cancers may be faster growing and more likely to spread. Grading was first introduced by Greenough in 1925 and modified by Bloom and Richardson in 1957, using three criteria; glandular formation, cell size and shape and proliferation. Currently the most commonly used grading system is the Nottingham histological grade [24]. Grade is moderately reproducible [25] but is nevertheless a prognostic factor used after adjustment for tumour size and lymph node involvement.

The most important predictive markers is the expression of two members of the steroid hormone receptor group, ER and PR. Expression of ER and PR is

associated with better survival and higher response rate to hormonal therapy. Women expressing both receptors have 80% better response, women expressing only ER have 30% and women who do not express either of the two receptors do not respond to oestrogen receptor modulators [26]. In addition, over expression of HER2/neu, a member of the HER family receptor tyrosine kinases, is seen in 30% of breast cancer cases. Breast tumours expressing HER2/neu are associated with more aggressive tumour characteristics and poor survival rates. Breast cancer cases over-expressing HER2/neu can be treated with trastuzumab [27].

The last few years a huge effort has been given to discover expression profiles that could eventually prove useful to predict the likelihood of breast cancer recurrence and response to treatment. The following four gene expression assays Oncotype DX (evaluates the likelihood of breast cancer recurrence and assesses the benefit of adjuvant chemotherapy), MammaPrint test (to predict the likelihood of breast cancer recurrence within five to 10 years), Rotterdam Signature (predict risk of breast cancer recurrence), and the Breast Cancer Gene Expression Ratio (to help predict recurrence of breast cancer) are under investigation and for the first two expression profiles trials are underway to confirm their clinical value [19].

2.1.4 Predisposing Factors

Epidemiological studies have suggested that sex, age, history of benign breast disease, breast cancer family history (particularly in a first-degree relative, such as mother and sister); genetic and epigenetic alterations, hormonal and reproductive factors, indicative of oestrogen exposure, such as age at first period, age at first pregnancy, parity, age at menopause, use of hormonal drugs and high

endogenous sex steroid hormone levels and behavioural and lifestyle factors such as diet, weight and alcohol intake affect breast cancer risk [28]. Hormonal and epigenetic factors which are the main subjects of this thesis are discussed in more depth in sections 2.2 and 2.3. Table 2-1 summarises predisposing factors for breast cancer.

Table 2-1: Predisposing factors for breast cancer.

BREAST CANCER RISK FACTORS			
Parameter	**Low Risk**	**High Risk**	**RR**
Sex	male	female	150.0
Age	Young	Old	>10
Family history	No	Yes	2.6
BRCA1 mutation	No	Yes	15
DNA methylation changes in tumour	No	Yes	1.4-5.3
History of benign condition	No	Yes	4.0-5.0
Age at menarche	>14	<12	1.5
Age at first birth	<20	>30	1.9 - 3.5
Age at ovarectomy	<35	no	3.0
Age at menopause	<45	>55	2.0
BMI (postmenopausal)	<22.9	>30.7	1.6
HRT	never	current	1.2-1.4
Bone density	1st quintile	4th quintile	2.7-3.5
Breast density	10%	>75%	4.6
Serum Oestradiol	1st quintile	4th quintile	1.8-2.4
Weight gain	Low	High	1.2-2.3
Height	Low	High	1.3-1.9
Radiation	No	Yes	1.6-5.2
Alcohol	No	Yes	1.4
Smoking	No	Yes	1.13-1.50

BMI=body mass index; HRT= hormone replacement therapy

(Adopted from Veronesi et al, 2005 [17] and further expanded based on the references used in sections 2.1-3)

Sex

Females are at a higher risk compared to males. Male breast cancer is very rare but this risk appears to be rising [11]. This difference could mainly be explained based on the different exposure to hormones between the two sexes.

Age

Age is one of the most important risk factor for breast cancer. The risk of getting breast cancer increases steadily with increasing age. Before the age of 25 breast cancer is rare and it tends to be more aggressive when it occurs. At the age of 30 there is a sharp increase with more than 80% of cases in women being diagnosed over 50 and with the greatest rate of increase prior to the menopause supporting an association with hormonal status [11]. Increased risk due to aging may be partly due to genetic [29] and epigenetic changes [30].

Perinatal

Breast cancer risk has been long hypothesised to be influenced by early-life exposures, or intrauterine events and conditions. This field started gaining ground after Trichopoulos *et al*, published a report about the intrauterine origins of breast cancer [31]. In two recent meta-analysis studies, it was shown that high birth weight as well as birth length are associated with an increased breast cancer risk. Pre-eclampsia, eclampsia and twin membership were associated with a reduced risk of breast cancer for both the mother and the offspring. Being breast-fed, gestational age and maternal diethylstilbestrol treatment did not seem to be associated with breast cancer risk in the offspring. Potential underlying mechanisms could include high levels of maternal endogenous sex and growth hormones, germ-cell mutations and other genetic-epigenetic events [32, 33].

Family history

Breast cancer family history among first or second degree relatives is one of the most important risk factors [34-36]. The relative risk of breast cancer is doubled for women who have a first-degree relative with breast cancer and risk is even higher if the relative is diagnosed before the age of 50 or when the woman has more than one relative affected by the disease [37].

Genetic changes

Breast cancer, like other tumours, develops and progresses through an accumulation of genetic alterations such as mutations and/or chromosomal alternations by activating oncogenes and/or inactivating tumour suppressor genes. In the 1990s, research into the genetic basis of breast cancer led to the identification of ‘'high risk'’ breast cancer susceptibility genes breast cancer 1 (*BRCA1*) [38] and 2 (*BRCA2*) [39]. Women carrying mutations on *BRCA1* and *BRCA2*, have a life-time breast cancer risk of 45-65% [40]. At present these are the most well described genes that increase not only breast cancer risk but also ovarian.

Although 10% of the breast cancer cases have been shown to be due to highly penetrant inherited susceptibility genes, it has been suggested that 27% of the breast cancer risk could also be attributed to other hereditary genetic factors [41]. Based on that hypothesis several studies have been carried out to identify novel hereditary genes. Genes that have been reported to be associated with breast cancer are *ATM* (Ataxia-Telengiectasia), *CHEK2* (Li-Fraumeni-like Syndrome), *STK11/LKB1* (Peutz-Jegher Syndrome), *PTEN* (Cowden Syndrome), TP53 (Li-Fraumeni); which also lead to rare conditions, *BRIP1* (*BRCA1* interacting protein

C-terminal helicase-1) and *PALB2* (the partner and localizer of *BRCA2* – PALB2 protein interacts with the protein produced from the BRCA2 gene. These two proteins work together to mend broken strands of DNA, which prevents cells from accumulating genetic damage that can trigger them to divide uncontrollably). However, these variants combined explain less than 1% of the genetic risk [42].

In the last few years a huge emphasis has been given to the identification of genes that appear to have a small effect on the risk (low penetrance) being either protective or additive, depending on the nature of the genes and the alleles involved [43]. Based on these studies it has been reported that a certain *CYP17* genotype (A2 allele) is a risk factor for breast cancer [44] being associated with elevated levels of sex steroid hormones, oestrogens and androgens, in both pre- and post- menopausal women [45, 46]. However, other studies showed conflicting results [47, 48].

With the introduction of whole genome-wide association studies 13 novel susceptibility loci have been identified with a very high statistical significance in populations of European ancestry [49]. These loci include: trinucleotide repeat containing 9 (*TNRC9*), mitogen-activated protein kinase 3 K1 (*MAP3K1*), TOX high mobility group box family member 3 (*TOX3*), the 8q24, lymphocyte-specific protein 1 (*LSP1*) and fibroblast growth receptor 2 (*FGFR2*) [49], with the latter being further validated and confirmed by another study [50]. In 2008, two studies by Garcia-Closas and her colleagues showed that some of these loci are related with clinically important characteristics such as ER and PR expression, grade and node involvement [51, 52]. In addition, non genomic regions on 2q23 and 5p12 were further identified and were shown to increase ER-positive breast cancer risk [53, 54].

More recently, a study identified five new susceptibility loci, on chromosomes 9, 10 and 11 and in 6q25.1 and LSP1 regions that showed more significant association with risk than those reported previously when familial breast cancer cases [55] were investigated. These novel loci could eventually be useful for population screening [42] and for improving diagnostic methods.

History of Benign Conditions

Studies have shown that history of benign breast diseases, including proliferative benign diseases with or without atypia, is a significant breast cancer risk factor [56]. Most benign conditions are non proliferative (cysts and fibroadenomas). Women having such a condition are not in a high risk to develop breast cancer [57, 58]. Women having a proliferative benign condition (hyperplasia, papilloma, radial scar) seem to be at 1.5-1.9 fold increased risk of breast cancer and women having atypical hyperplasia, either lobular or ductal types, seem to be at 4-6 fold increased risk to develop breast cancer. Nearly 40% of women with a family history and atypical hyperplasia subsequently develop breast cancer [59-61].

Breast density

Breast density reflects the breast tissue composition and is associated with collagen, epithelial and non-epithelial cells. A case-control screening population study demonstrated that women with higher breast density compared to those with lower were nearly five times more likely to develop breast cancer [62]. Still it is unknown whether breast density changes over time influence breast cancer risk [63], but it has been suggested that women leaving in urban areas have denser breasts compared to women leaving outside the city increasing the attention to screening in these regions [64].

2.1.5 Detection and Screening

Currently available detection techniques for breast cancer screening include imaging tools (mammography, ultrasonography, magnetic resonance imaging (MRI) and positron emission tomography and analysis of tumour markers. Physical examination is also another important detection method since almost 33% of women developing breast cancer are not identified by imaging tools.

Imaging

Mammography is one of the most important diagnostic tools in postmenopausal women but it is not considered a good method to detect breast cancer in dense breasts and recognise certain ILCs, Paget's disease, inflammatory and small peripheral carcinomas [65, 66]. In UK, the National Health System (NHS) by the Breast Screening Programme offers free mammogram to all women aged 50 to 70. It has been shown that mortality rates have been falling after the introduction of mammography screening programs [67-70] but there has been a lot of controversy in the last few years as to its true value [2, 3, 71]. Ultrasonography has been suggested to be an effective procedure to diagnose small tumours in women with dense breast [72]. MRI was shown to be a highly sensitive technique for screening high-risk patients who are younger than 50 years [73]. The UK MRI Breast Cancer Screening study reported that this method can detect twice the number of women compared to mammography [74]. However, even though MRI has good diagnostic accuracy, the rate of false-positive cases is still high and its findings cannot be used as the only source of information for surgery decision [75]. Position emission tomography is presently used to discover metastatic tumours in distant organs and is able to assess the status of axillary nodes in pre-operative staging processes [76].

Tumour Markers

Tumour markers are certain characteristics of the tumour that differ from the normal tissue and can be visible and/or measurable effects of tumourigenesis [10]. The efficacy of a biomarker is determined by its sensitivity and specificity. The clinical sensitivity of a biomarker refers to the proportion of subjects with confirmed disease who test positive, whereas its specificity refers to the proportion of healthy control subjects who test negative [77]. In order to have an effective detection biomarker it needs to be: a) non-invasive or minimally invasive, b) be measured using a small amount of specimen c) site-specific, being able to rule out non-cancerous events in the same organ or tissue, d) highly specific in order to limit false-positive results, e) simple to perform with low cost and f) observer-independent [10]. An example of a biomarker that could prove valuable for ovarian cancer screening is cancer antigen 125 (CA125). The largest randomised control trial (UKCTOCS) [78] coordinated by our group is underway to establish the biomarker's potential in clinical setting in combination with sequential ultrasonography.

Taking into consideration that any biological molecule can be used to distinguish normal from abnormal samples; different analytes have been examined as possible biomarkers of breast cancer. The biological molecules include: lipids, carbohydrates, polyamines, proteins (**proteomics**), RNA (**genomics**) and DNA (**genetics and epigenetics**) being examined in various sources, such as tissue, serum, plasma, nipple aspirate fluid and ductal lavage fluid [10].

Proteomics detect the function of expressed genes through biochemical analysis of proteins providing a dynamic and accurate reflection of both the intrinsic

genetic programme of the cell and the impact of its immediate environment. Therefore, given that proteomics can provide link between gene sequence and cellular physiology it is suggested that the proteome could complement the genome for biomarker discovery. Recent advances in mass spectrometry, such as matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry (MALDI-TOF MS) and its variant surface-enhanced laser desorption/ionisation (SELDI-) TOF MS, have enabled high-throughput proteome analysis [79]. Several proteins have been identified to have significant breast cancer diagnostic, prognostic or predictive value. The most widely used tumour markers of the breast include: CA15-3 and CA27.29 [79, 80]. Both markers are also known as the extracellular mucin 1 (MUC1) protein (antigens for MUC1). MUC1 has been shown to be over-expressed and aberrantly glycosylated in many cancers. Recently, autoantibodies generated to aberrant O-glycoforms of MUC1 were evaluated to see whether they would serve as diagnostic biomarkers. Cancer-associated immunoglobulin G autoantibodies in serum of breast cancer patients against different aberrant O-glycopeptide epitopes derived from MUC1 were detected representing sensitive biomarkers for early detection of breast cancer [81]. Regardless of the voluminous data on proteins, their value as markers in large, prospective, clinical studies still needs to be shown. To date the only cancer biomarker being recognised as possibly effective for screening is the prostate-specific antigen for prostate cancer with ongoing trials to prove its value [82].

Genomics are defined as the measurement of gene expression from available sequence information. The expression profile represents the function and phenotype of a cell and is called transcriptome. Several technologies including cDNA and oligonucleotide arrays and multiplex Polymerase Chain Reaction

(PCR) [83] have been developed to generate molecular signatures, that could eventually prove useful early detection biomarkers. Studies have suggested that mRNA in plasma as a tumour marker could facilitate the detection of cancer cases with high sensitivity [84]. Further studies are needed though since it is still unclear how stable RNA is in the bloodstream especially when high levels of serum ribonuclease have been detected in cancer patients [10].

Genetic and **molecular changes** causing genetic instability are early events occurring in carcinogenesis making them useful markers to detect breast cancer before the onset of symptoms and morphological changes. Such signatures have been studied using high throughput technologies and several markers are being extensively studied in animal models and in patients with established breast cancer. Signatures of interest include microsatellite instability, single nucleotide polymorphisms and **epigenetic changes** such as DNA methylation modifications. DNA methylation alterations occur at a high frequency, are reversible upon treatment with pharmacological agents, and arise at defined regions within a gene making them an attractive alternative for cancer detection and assessment [85]. Additionally, it is evident that body fluids can carry DNA methylation imprints demonstrating their possible diagnostic and predictive importance [86]. Various studies have reported the diagnostic potential of circulating tumour related methylated DNA in serum for cancer detection. Examples include the colon cancer specific methylation of septin 9 (*SEPT9*) which was shown to produce a specificity of 95% and a sensitivity of 52% when plasma samples were analysed and a panel of tumour suppressor genes (*APC*, *RASSF1A* and *p14*) that showed hypermethylation in bladder cancer with 87% sensitivity and 100% specificity [87].

2.1.6 Diagnosis

Breast diagnosis employs cytology by means of fine-needle aspiration (FNA) or histology by means of core needle biopsy or excisional biopsy. FNA has a high diagnostic accuracy, with 10-15% false negative rate. A rare false positive rate has been reported in association with ductal or lobular hyperplasia. In this technique a needle is inserted into the mass to extract cells which are stained and observed under the microscope to investigate any abnormal cell morphology. Core needle biopsy is a less invasive method in comparison to an open biopsy and utilises a needle to obtain the specimen. Abnormal architecture, invasion and specific tumour markers such as ER, PR and HER-2/new can be evaluated in these tissue samples providing more information than FNA. Open biopsy which can be guided by image tools (mammogram or ultrasound), is rarely used and is only provided when inadequate sample from a core needle biopsy is obtained or the pathologic results are equivocal [88].

2.1.7 Treatment

The widespread adoption of effective treatments has resulted in a rise of the 5 year mortality to 75% [11]. The current breast cancer treatment involves multimodality therapy including: surgery, radiotherapy, chemotherapy, endocrine and molecular therapy, with systemic treatment being given before (neo-adjuvant) or after (adjuvant) surgery [89-92].

Surgery has always been the primary treatment for breast cancer. Depending on the tumour size, breast conserving surgeries, such as lumpectomy and wide local excision (WLE), followed by breast irradiation is followed. Axillary lymph node clearance (ANC) is an important surgical procedure for invasive breast cancers and staging the tumour.

Women who are at high risk of recurrence are treated with radiotherapy. In a meta-analysis of randomised trials it was shown that radiotherapy reduces mortality rates [93]. The impact on mortality though has been debated as radiotherapy has also been shown to increase the risk of cardiovascular events [94]. Intraoperative radiotherapy has been suggested to be the answer to the problems that surround conventional radiotherapy. A multinational clinical trial in UK, TARGIT (Targeted Intraoperative Radiotherapy Treatment), recently published a report suggesting targeted intra-operative radiotherapy, using conventional external beam, could be an attractive alternative to conventional postoperative radiotherapy for the treatment of early stage breast cancer [95-97].

Endocrine therapy is the second key treatment of adjuvant therapy and for more than twenty years tamoxifen is the most commonly used drug [98]. Tamoxifen is a member of selective ER modulators [99] and it is an anti-oestrogen that blocks the binding of oestrogen to its receptor. Therefore, only hormone dependent tumours are treated with the drug [100]. Tamoxifen is effective in both pre- and post-menopausal women, and when it is given after chemotherapy rather than at the same time [90], the standard therapy duration is 5 years. The main disadvantages are that it has antagonistic and agonistic functions on other organs such as the endometrium and bone [101] and cause an increased risk of thromboembolism. After long administration breast cancer patients can become resistant. In the last few years, clinical trials have shown that aromatase inhibitors are also suitable for breast cancer treatment [102]. These drugs inhibit the conversion of sex steroid hormones to oestrogens. They include the non-steroidal anastrozole and letrozole and the steroidal compound exemestane. They have been shown to be effective only in postmenopausal women and to be superior to tamoxifen alone

(Arimidex, Tamoxifen, Alone or in Combination trial also known as ATAC trial) by improving the disease and metastatic free survival [103]. Patients are given aromatase inhibitors after two or three years of tamoxifen, as it was shown that could improve the long term survival and reduce breast cancer recurrence [104]. In addition, the value of ovarian suppressors is investigated with trials undergoing and alredy available data suggesting no added effect of ovarian ablation or suppression on the relapse-free survival or overall survival of premenopausal women who were treated for early-stage breast cancer. However, the role of ovarian ablation or suppression with ER-positive tumours requires further investigation [105].

Chemotherapy is usually selected for women with high risk of metastatic disease and poor prognosis [106] given as an adjuvant treatment after surgery to increase the chance of long-term disease free survival and as neo-adjuvant treatment to reduce the size of the tumour before surgery. The most commonly used chemotherapeutic agents are the anthracyclines (doxorubicin and epirubicin). These drugs have been shown to be more effective than the traditionally used cyclophosphamide, methotrexate and fluorouracil chemotherapeutic agents [91]. Furthermore, evidence indicates that addition of a newer class of chemotherapeutic agents, the taxanes (paclitaxel and docetaxel), may improve survival in high risk patients [106].

In addition, in the last decade there is an increasing interest in antibody treatment, after it was shown that growth factor receptors are correlated with poor disease-free survival and resistance to endocrine therapy and chemotherapy, inhibiting growth factor activity [107]. Drug development strategies include anti-

receptor antibodies and tyrosine kinase inhibitors which target the epidermal growth factor receptor (*EGFR*) and *HER-2/neu*. The most well know is Herceptin, a humanized monoclonal antibody that is directed against the external domain of the *HER2* receptor [108]. A recent study has demonstrated that yearly administration of Herceptin during or after chemotherapy can reduce recurrence risk by 50% [109]. Another drug is the tyrosine kinase inhibitor lapatinib [110]. The agent targets *EGFR* and *HER2* receptors and prevents tumour growth by inhibiting intracellular tyrosine kinase activity by binding to the inactive form of the receptor and dissociating at a slow rate having a better effect on the target site [111].

Epigenetic therapy has also drawn attention as epigenetic changes can be reversed by use of small molecule inhibitors with restoration of the affected epigenome [112]. The discovery of 5-azacytidine (5-aza-CR) (vidaza) and 5-aza-2'-deoxycytidine (5-aza-CdR) (decitabine) agents were the first steps towards epigenetic treatment. These agents can incorporate into the DNA of rapidly dividing tumour cells and reverse the action of enzymes that are responsible for DNA methylation alternations (DNA methylation transferases – DNMTS) [113] causing DNA demethylation and reactivation of methylated silenced tumour suppressor and other cancer-related genes [114]. The use of these drugs though has been hindered by their cytotoxic side effects which result from their incorporation into the DNA [115]. An alternative approach is the use of non-nucleoside DNMT inhibitors, such as MG98, SGI-1027 and RG108 which exert their effect without being incorporated into the DNA [115]. In order to better understand the full potential of epigenetic therapy in the clinical setting, more knowledge is required of the molecular action of these agents [116].

2.1.8 Risk Prediction and Prevention

The last decade there is growing interest in trying to stratify women into groups based on different levels of breast cancer risk. Currently, even though mammography screening has been suggested to decrease mortality from breast cancer [117] - with its true value being debated [71] - it does not reduce the number of women who develop the disease [118]. Tools such as the Gail, Claus and Golditz models have been developed to calculate a woman's absolute risk of breast cancer. These models help clinicians to identify women whose breast cancer risk is increased based on their epidemiological profile. In addition, there are models such as the Tyrer-Cuzick model [119] and the BRCAPRO program [120] that calculate the likelihood of a women with a breast cancer family history to have a *BRCA* mutation in addition to breast cancer risk. The last few years it has become increasingly recognised that breast cancer is not a homogeneous disease with the hormone-receptor status defining important clinical and aetiological differences. Based on that observation a model by Rosner and Colditz has been developed which separately estimates the risk of hormone-receptor-positive and hormone-receptor-negative breast cancers [121].

At present, the only model that is widely used in family clinics is the Gail model [4, 5]. The model includes the following risk factors: age, race, age at menarche, age at first live birth, number of first degree relatives with breast cancer, number of previous breast biopsy examinations and presence of atypical hyperplasia. In recent attempts to improve Gail's model performance breast density was included as an additional parameter and the test's concordance statistics were brought up from 59% to 66% [7, 8]. In a different study, single nucleotide polymorphisms were included in the model and the test's concordance statistics were brought up to

61.8% [122]. When oestradiol was added in the Golditz model the test's concordance statistics were brought up to 64.5% [123]. A study has also been performed to examine performance of the Gail model for estimating invasive breast cancer risk by receptor status in postmenopausal women. The data showed that the model's discriminatory performance was better for ER-positive breast cancer risk [124]. Attempts to increase model's performance is significant as identification of women at high risk could eventually lead to improved overall survival rates of breast cancer patients through prevention and more intensive screening.

The current preventative options for breast cancer include changes in lifestyle, chemoprevention and secondary prevention by detection of cancer by regular surveillance either through palpation of the breasts and/or regular participation in mammography and MRI. It has been shown that prophylactic surgery, either bilateral mastectomy or oophorectomy, is the most cost effective means to reduce breast cancer risk and it is recommended to women with high risk of hereditary disease [125]. Chemoprevention is another approach and it has been shown that there is a significant reduction in breast cancer incidence in women taking tamoxifen [126] and an even larger reduction in women taking raloxifene [127]. The true value of such prevention strategies has been argued though with studies showing conflicting results [125]. Moreover, tamoxifen and raloxifene have well known side effects such as endometrial cancer (tamoxifen) and thromboembolism (raloxifene) causing caution regarding their use [127]. In addition, neither drug has been shown to be of value in preventing ER-negative breast cancers. Therefore their true benefit must be considered carefully before prescribing them to healthy women [127]. The role of aromatase inhibitors in breast

cancer prevention is still under investigation [128]. Careful consideration between the risks and the benefits of the different prevention strategies is important in order to have the best impact in women's life.

2.2 Hormones and their Receptors

In 1905, Ernest Starling a professor of physiology at University College London, UK, was the first to use the word 'hormone' in one of his lectures. Starling defined the word, derived from the Greek word ''orme'' meaning ''to arouse or excite'', as ''the chemical messengers which speeding from cell to cell along the blood stream, may coordinate the activities and growth of different parts of the body''. Later on, it was shown that cells respond to a hormone when they express a specific receptor for that hormone. The hormone binds to the receptor proteins that result in the activation of a signal transduction mechanism that ultimately leads to cell type-specific responses [129]. Several hormones have been studied to examine their association with breast cancer risk. In this thesis, sex steroid hormones, their receptors and gonadotrophins are discussed in more depth in the following sections.

2.2.1 Sex steroid hormones and their receptors

Sex steroid hormones include oestrogens, androgens and progestogenes. They are hormones that interact with the vertebrate ER, AR or PR, respectively also known as nuclear receptors. Natural sex steroids are made by the gonads; in females by ovaries and males by testes, and in both genders by the adrenal glands and by conversion from other sex steroids in tissues such as liver or fat catalysed by specific enzymes. They are lipophilic molecules derived from cholesterol which is the main precursor of all sex steroid hormones. Circulating

cholesterol carried by low-density lipids can either be used immediately for hormone synthesis or be stored as cholesterol esters [130].

The first and rate limiting step in sex steroid biosynthesis is the cleavage of cholesterol's side chain to produce pregnenolone, a reaction catalysed by the cytochrome P450 side-chain-cleavage enzyme ($P450_{scc}$). Pregnenolone is then converted to either 17α-OH pregnenolone by $P450_{17}$ an enzyme that also catalyses the removal of C20-C21 side chain of 17α-OH pregnenolone converting it to DHEA or to progesterone by 3β-hydroxysteroid oxidoreductase-$\Delta^{4,5}$-isomarase enzyme complex which then converts DHEA to androstenedione. Through the action of 17-oxo-reductase androstenedione is converted to testosterone. Through aromatisation then androstenedione is converted to oestrone and testosterone to oestradiol (Figure 2.1). Alternatively, DHEA is converted to androstenediol by 17-oxo-reductase and then to testosterone by 3β-hydroxysteroid oxidoreductase-$\Delta^{4,\ 5}$-isomarase enzyme complex. The adrenal glands secrete large amounts of DHEA and its sulphate metabolite DHEA-S serving as precursors of both androgens and oestrogens in peripheral tissues. In the ovary where androgens and oestrogens are produced there are high levels of $P450_{17}$ and upon activation of LH receptor in theca cells cholesterol is converted to testosterone and after stimulation by FSH oestradiol is produced in granulosa cells through aromatisation of testosterone [130]. In blood the main fraction of different hormones are bound to albumin or serum SHBG and thereby inactivated. The unbound fractions may diffuse through cell membranes due to their lipophilic properties and bind to the intra-nuclear, mitochondrial or other intracellular sex steroid receptors.

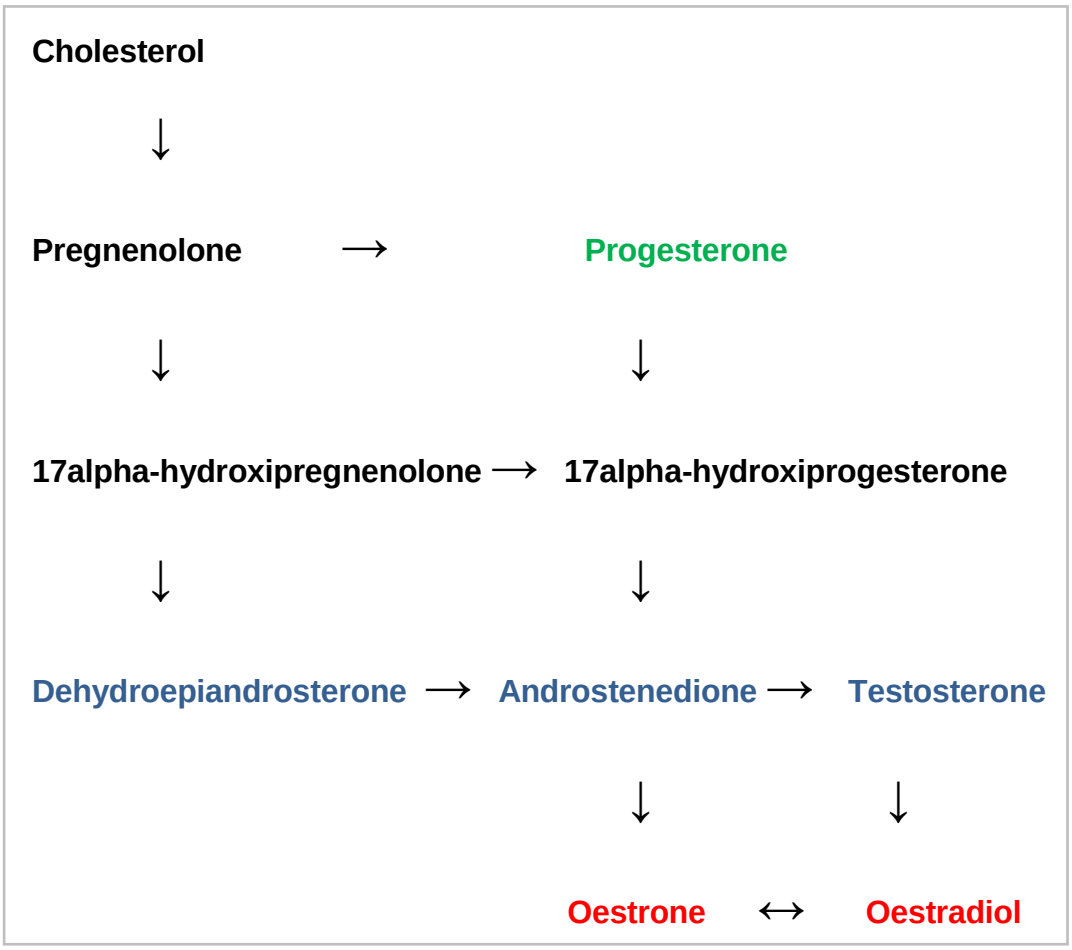


Figure 2.1: Summary of sex steroid biosynthesis. (Progesterone is in green, androgens in blue and oestrogens in red)

Sex steroid receptors belong to the nuclear receptor family. Hormone receptors belonging to this group have a common functional structure which is composed of independent domains: the NH_2-terminal domain which contains the activation function (AF) 1, the binding domain (DBD) which contains two zinc fingers binding to specific sequences of DNA known as hormone response elements (HRE), the ligand binding domain (LBD) with the hormone dependent AF2 domain, the hinge domain which connects DBD with LBD and is thought to allow the receptor to adopt to several different conformations and the C-terminal domain. The DBD

and LBD are highly conserved to provide specificity for the target hormones and genes. The NH_2-terminal, the hinge and C-terminal domains are highly variable between the different nuclear receptors. While these are the main domains shared by all nuclear receptors there are sites which are responsible for the binding and interaction of multiple accessory proteins that assist and modulate the functions of the receptors [131, 132].

In the absence of a ligand sex steroid hormone receptors are bound to molecular chaperones, such as hsp90, or interact with co-repressor proteins that repress the transcription of target genes. The receptors function once the ligand binds resulting in dissociation from the chaperone complex or the co-repressor proteins and interact with co-activator proteins mediating the assembly of the basal transcription machinery components and subsequent transcriptional activation of target genes. Upon that binding the activated receptor molecule, which is usually found in the cytoplasm of the cell, travels across the membrane into the nucleus where it gets homodimerised or heterodimerised with another isoform of the receptor or a different receptor and binds to the specific HRE. Once bound the receptor complex alters the transcription of the target genes either up or down regulating them [133] (Figure 2-2).

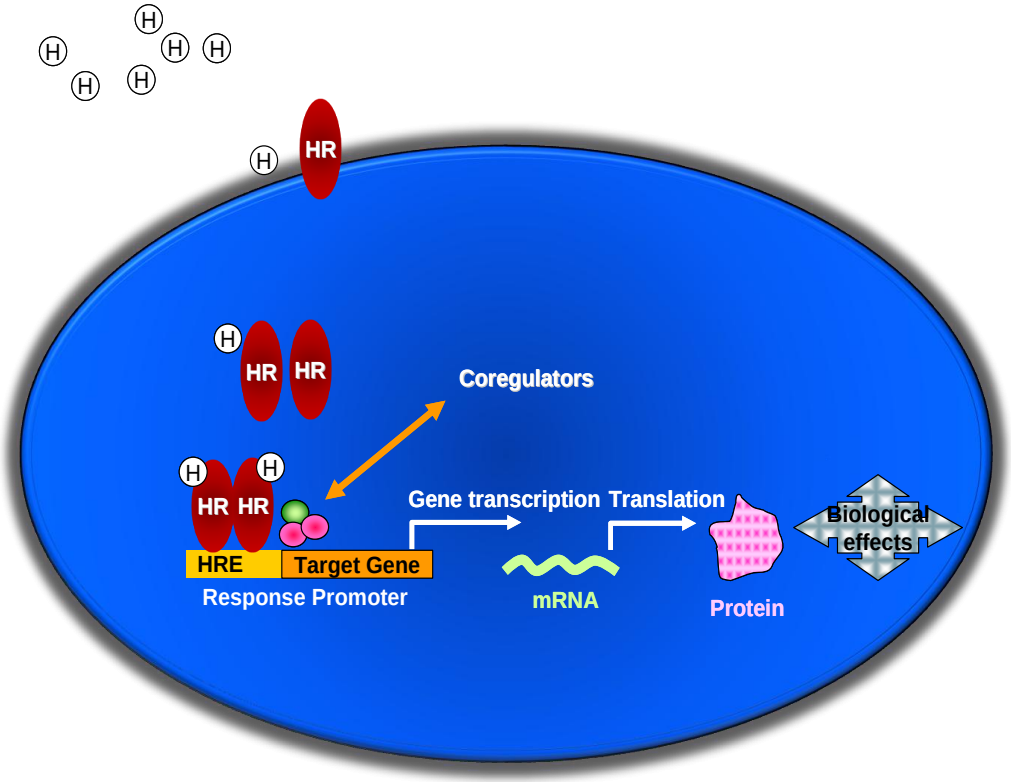


Figure 2-2: Activation of sex steroid hormone receptors. Hormone (H) enters the cell by passive diffusion; the hormone binds to the intracellular receptor (HR) which is in its inactivated state located either in the cytoplasm or the nucleus. Upon binding, sex steroid hormone receptor forms a dimer which then binds to specific sequence of DNA known as Hormone Response Element (HRE) leading to activation of the transcription processes and synthesis of specific messenger RNA and protein production.

Response element bound steroid receptors are known to additionally bind to chromatin remodeling complexes that cause rearrangement of the chromosome structure. Up-regulation of genes by the steroid receptors is usually facilitated by altering the chromosome structure so that the promoter and DNA polymerase binding sequences of the gene are exposed to the enzymes that initiate transcription. Steroid receptor mediated repression of gene expression is achieved through different mechanisms. In such events the HRE overlaps the TATA box, the DNA sequence that acts as a recognition site for the TATA binding protein (TBP) which is one of the proteins responsible for the initiation of transcription. When the activated steroid receptor binds to the HRE it displaces the bound TBP preventing the initiation of transcription. The second mechanism responsible for a steroid receptor gene expression switch is when the activated receptor is influenced by proteins bound at another site in the promoter of the gene. If the second binding site is bound by a homodimer or another protein then transcription is increased but if the site is bound by a heterodimer the transcription is stopped [133, 134].

The direct effect of nuclear receptors on gene regulation takes hours as a functional effect is seen in cells after a large number of events occur between nuclear receptor activation and protein production. However, it has been observed that some effects in the presence of the sex steroid hormones occur within minutes in contrast to the time consuming mechanism of nuclear receptor action. Therefore, a different mechanism has been suggested. This is known as non-genomic effect of the nuclear receptors. Data has shown that sex steroid hormones can interact with their sex steroid receptors at the plasma membrane and activate intracellular signaling cascades without the involvement of HREs [133].

The specific effects of sex steroid hormone receptors can also be modulated by the binding of specific co-factors to the activated receptor, splicing of mRNA before translation and modification to the translated protein. The different co-factors present within the cell can alter the effect of hormone stimulation by changing the affinity of the receptor for HRE binding. In addition, alternative splicing into different receptor isoforms is one of the major mechanisms to modulate the effect of hormone stimulation. The different isoforms have an altered affinity for the receptor's ligand, other receptors, hormone binding elements and receptor co-factors. The most common change in an isoform that has reduced affinity to the receptor ligand, results in an altered response to hormonal stimulation [133].

Post-translational modifications have also been shown to be responsible for differentiated response of the cells to hormone stimulation either directly or through the binding of specific co-factors to receptor complexes. A post-translational modification is the addition of a small molecule to the protein at specific sites by enzymes or enzyme complexes that modify the behaviour of the protein. Such modifications include phosphorylation, acetylation, methylation, sumoylation, sulphation and ubiquitination. For example, phosphorylation is known to increase the activity of the ligand bound receptor in contrast to ubiquitination which is responsible for removing the tagged protein to the proteosome for degradation resulting in reduction of the hormone stimulation [133].

Oestrogens

The term oestrogenes come from the Greek word ‘‘oistros’’ which refers to the phase when women can become pregnant. The major naturally occurring

oestrogens are 17-β oestradiol (E_2) and two metabolites oestrone (E_1) and oestriol (E_3). E_2 is the predominate form in non-pregnant females, E_1 is produced during menopause and E_3 is the primary oestrogen during pregnancy. Oestrogens are produced primarily by developing follicles in the ovaries and the corpus luteum in women in reproductive phase of their life. Oestrogens are also produced in smaller amounts by other tissues such as the liver, adrenal glands and fat cells. These secondary sources of oestrogens are especially important in postmenopausal women in whom these are the main sources of oestrogen production [130].

Oestrogen Receptor

The cDNA encoding an ER protein was first cloned and described in 1973 [135]. The name was changed to ER-α when a second form of the receptor, ER-β, was discovered in 1996 [136]. ER-α is slighter bigger than ER-β encoding a protein that is 595 amino acids in comparison to ER-β which encodes a protein that is 530 amino acids long. The two different forms of ER are known to be each encoded by two separate genes, *ESR1* and *ESR2* for ER-α and ER-β, respectively. *ESR1* is located at chromosome 6q25.1 and *ESR2* is located in chromosome 14q22-24 [137, 138]. *ESR1* and *ESR2* genes show significant overall sequence homology. In the DBD there is only one base pair difference with 97% homology and 59% amino acid identity in the LBD. This homology is lost at the end of the genes, in the trans-activation factor domains, modulating their receptor functions that results in different effects [139, 140]. In summary, despite the differences in structure between the two receptors their expression is regulated by the same factors and have similar binding affinity to free oestradiol but their molecular and transcriptional activity and their tissue localisation are different. ER-α is

expressed in endometrium, ovarian stromal cells, hypothalamus and breast cancer cells, ER- β is expressed in kidney, brain, heart, lungs, intestinal mucosa and endothelial cells [141].

Androgens

Androgens, including testosterone and 5α-dihydrotestosterone (DHT) control the development, differentiation and function of male reproductive and accessory sex tissues such as prostate. In females, androgens are mainly the precursors of all oestrogens. There are also the adrenal androgens produced by the adrenal cortex that function as weak steroids or steroid precursors including DHEA, DHEAS, and androstenedione. Testosterone is the most important circulating androgen in both men and women. Its effect is mediated directly via AR binding or after peripheral aromatisation to oestradiol via the ER or through 5-a reduction in the form of DHT [130].

Androgen Receptor

The AR gene is located at Xq11-12 and is 90kb long with eight exons. There are two characterised forms of AR -A and -B. The second form, which is also the most predominant one, is a 110-114 kDa protein of 910-919 amino acids. The first is a smaller protein, 87 kDa, of 720-729 amino acids in length that makes up only about 4-26% of the detectible androgen receptors located in varying tissues and it is believed to be inert with its exact function being unknown [142-144].

Progestogenes

Progesterone is the naturally occurring progestogen also known as hormone of pregnancy. It is synthesised in the ovary, adrenal gland and during pregnancy by

the placenta. It serves as a precursor for the production of oestrogens, androgens and adrenocortical steroids. It is responsible for the preparation of the endometrium for implantation, keeping the myometrium quiescent until parturition and affecting the female immune system to accept the foetus [130].

Progesterone Receptor

PR gene is located at 11q22-23 chromosome and contains eight exons and is about 90kb [145]. There are two forms of the receptor PR-α and PR-β both of them deriving from the same gene but being activated by different promoters. PR-α is a 98 kDa protein and PR-β is larger than the alpha isoform at 116 kDa [146]. Both isoforms have different effects on different tissues because of their different structure and accessory molecules helping to initiate or block transcription [147].

2.2.2 Gonadotrophins and their receptors

Gonadotrophins are protein hormones that are secreted by gonadotrope cells from the pituitary gland. The two principal gonadotrophins are LH and FSH. Chorionic gonadotrophin (hCG) belongs also in the group and is produced by the placenta during pregnancy.

Luteinising hormone and Follicle stimulating hormone

LH and FSH are heterodimers consisting of two peptide chains, alpha and beta. LH and FSH share nearly identical alpha chains (LH and FSH - 92 amino acids) whereas the beta chain is different (LH-121 amino acids; FSH-118 amino acids). The beta chain is the one responsible for providing specificity for receptor interaction. Their release is controlled by gonadotrophin-releasing hormone from the hypothalamus [130].

Luteinising hormone receptor and Follicle stimulating hormone receptor

The luteinising hormone receptor (LHR) and follicle stimulating hormone receptor (FSHR) are transmembrane G protein-coupled receptors. They are found in the ovary and the uterus. LHR interacts with LH and hCG and FSHR interacts with FSH. The gene of both receptors is located on chromosome 2p21. LHR consists of 70 kilo base pair (bp) and FSHR consists of 2,080 nucleotides. LH is 674 amino acids long and has a molecular mass of about 85-95 kDA and FSH is a 695 amino acids protein and has a molecular mass of about 76 kDa. Both receptors have the same structure consisting of an extracellular domain, a transmembrane domain and a C-terminal domain. Upon binding of the hormone a transduction signal takes place that activates the G protein. After attachment of the hormone on the receptor the cAMP system gets activated and shifts hormone's state from inactive to active [148].

2.2.3 Hormone detection methods

Hormones are generally measured with immunoassays. However other techniques and in particular bioassays could also play a significant role. Both assays relay on a comparison between responses produced in the assay system by the sample and those produced by the different concentrations of a reference sample. A calibration curve is generated with the reference preparation and the unknown concentration or bioactivity of the hormone in the sample can then be extrapolated from this. Immunoassays are based on the interaction between an antibody and its antigen and there are two different forms; competitive-binding assay and sandwich assay. In competitive-binding assay the antigen in the unknown sample competes with the labelled antigen to bind with antibodies and the amount of the labelled antigen bound to the antibody site is measured. The

response is inversely proportional to the concentration of antigen in the unknown sample because the greater the response the less antigen in the unknown sample is available to compete with the labelled antigen. In sandwich assay the antigen in the unknown sample is bound to the antibody site and then labelled antibody is bound to the antigen and the amount of labelled antibody on the site is measured. The response is directly proportional to the concentration of the antigen, in contrast to the competitive method because the labelled antibody cannot bind if the antigen is not present in the unknown sample [130].

Bioassays are an attractive alternative for hormone measurement. They are conducted to measure the effects of substances on a living organism and involve estimation of the concentration or potency of the substances by measurement of the biological response that produce. Receptor bioactivity assays have been of great interest as they allow evaluation of the overall hormonal effect instead of measuring single hormone compounds; therefore, it can be used to estimate circulating hormonal bioactivity. Their advantage is that they are based on a direct interaction between the binding ligand and the relevant bioactive site on the structure of a hormone and are characterised by all the positive features of the immunoassays in terms of sensitivity, precision and high sample capacity [130].

Another method to measure hormones is by mass spectrometry (MS). MS is an analytical technique that determines the elemental composition of a sample or molecule. Its principle consists of ionising chemical compounds to generate charged molecules or molecule fragments and measurement of their mass-to-charge ratio. Even though MS is a very sensitive technique, there are still some issues regarding its reproducibility [149]. Additionally, it is a very laborious technique

and requires a large amount of starting material making difficult to implement in clinical settings [150] and widely used [149].

During the last few years concerns have been raised regarding the reliability and validity of steroid sex hormone measurements in biologic specimens using immunoassays. Studies performed to evaluate the reproducibility of these assays have shown that there is considerable variation in results from different laboratories but measurements from a single laboratory are satisfactorily reproducible [151]. Comparison of immunoassays to MS has shown that both techniques yield similar estimates of most sex steroid hormones [151, 152].

2.2.4 Sex steroid hormonal changes and their receptors during normal development

The main role of sex steroid hormones is the induction of the primary and secondary sex characteristics and skeletal maturation. The last few years they have also drawn attention due to their profound regulatory effects on differentiation and growth and their function in a variety of tissues including brain, cardiovascular and adipose tissue.

Sex steroid hormones are responsible for signalling the development of the glandular breast tissue. All women throughout their life cycle experience changes in their breasts due to fluctuating sex steroid hormone levels during the menstrual cycle. The most obvious changes occur during foetal development, puberty, pregnancy and menopause. Human breast tissue fast develops in the sixth week of foetal life. After birth there is little development till puberty when in women, breasts begin to grow with the production of oestradiol. During this time, fat and

fibrous breast tissue becomes more elastic, the breast ducts begin to grow and this growth continues until menstruation begins. During pregnancy, a variety of breast changes occur. The blood vessels within the breast enlarge as surges of oestrogen stimulate the growth of the ducts and surges of progesterone cause the glandular tissue to expand. During menopause there is a huge decline of oestrogen and progesterone production affecting the structure of the breast. The breast shrinks, the glandular tissue gets replaced with fatty tissue and the fibrous-connective tissue loses its strength. Changes in circulating androgen levels with age have not been well documented due to lack of longitudinal studies [153] but is well known that testosterone levels are mainly maintained after menopause [154].

The members of the sex steroid receptor family have numerous functions and are associated with the control of cellular growth and differentiation in many tissues. In the human breast, ER-β is the predominant ER expressed in the luminal epithelial cells and the stromal compartment, in contrast to ER-α which is expressed in low levels only within the luminal cells. ER-α is the receptor which functions primarily as a mitogenic factor in the breast tissue while ER-β has been suggested to antagonise the functions of ER-α [155]. AR is important on the breast mainly due to the effects of androgens on the breast tissue. PR has two main actions in normal breast tissue; interferes with ER and slows the growth of breast epithelial tissue [147]. Since the physiology of breast is different before and after menopause, it is important to be taken into account in study designs. Therefore, studies that are affected by the different structure and biochemistry of the breast should separate women based on their menopausal status. For the purposes of this study, only postmenopausal women have been studied.

2.2.5 Gonadotrophins during normal development

LH and FSH are very important hormones during reproduction exerting their effects through their receptors. LH initiates steroidogenesis in the ovarian follicle, induces ovulation and maintains secretory functions of the corpus luteum. FSH stimulates the development of ovarian follicles and secretion of oestradiol. After menopause due to low oestrogen production LH and FSH levels rise [130].

2.2.6 Sex steroid hormones and their receptors in breast cancer

Sex steroid hormones have been long hypothesised to increase breast cancer risk. The connection between sex steroid hormones, and in particular oestrogens, with breast cancer was established more than 100 years ago by George Beatson, when he observed a drastic clinical response for breast cancer cases after removal of the ovaries suggesting that the ovaries have a control over the proliferation of the breast epithelium [156]. Later on, in the 1960s and 1970s, reproductive risk factors, which regulate sex steroid hormone availability in women, such as ages at menarche, first birth, parity and menopause were shown to be associated with breast cancer risk and based on these factors Malcolm Pike suggested an age-incidence model for breast cancer [157].

Further epidemiological studies supported the relationship that surrogates for lifetime sex steroid hormone exposure and exogenous sex steroid hormones being associated with breast cancer risk. Reports have shown high levels of oestrogens to be related with increased breast cancer risk. Oestrogens are associated with both the initiation and progression of breast cancer through two pathways; one that involves the binding of oestradiol to ER-α stimulating cell proliferation leading to an increased number of DNA replications followed by an

increased number of replication errors and the second one which is a result from the formation of genotoxic metabolites of oestradiol, which can bind to DNA causing depurination that eventually leads to mutations [158, 159]. While the majority of breast tumours respond to oestrogens, some lose expression of the ERs either because the gene becomes disabled or because the receptors are spliced or mutated. These tumours are not responsive to hormonal treatment.

In breast cancer the predominant receptor is ER-α being expressed in 70% of breast cancer cases. After the discovery of ER-β the complexity of oestrogen signalling in the breast was noted. ER-β has been shown to be expressed only in advanced breast cancer and in low levels [160] and has been suggested to act as a negative modulator of ER-α, changing its transcriptional activation by altering the recruitment of c-Fos and c-Jun to oestrogen-responsive promoters [161]. Even though, ER-α is a well known and used as a prognostic and hormonal treatment predictive factor, only the last few years studies supported the hypothesis that increased expression of ER-β is associated with an increased likelihood of response to endocrine therapy [160]. Recently, our group was the first to provide evidence that serum ER-α and -β bioactivity are independently associated with breast cancer [162]. It was further shown that sex steroids through their receptors can modify methylation patterns in the DNA of cells that are not directly related to the target organ and this can be used to assess breast cancer risk [163].

Although oestrogens and ER in breast cancer have been extensively studied the role of androgens and AR in breast cancer have been less investigated and remain poorly understood. There is evidence through retrospective and prospective studies that increased levels of androgens are significantly

associated with higher breast cancer risk. I*n vitro* studies have shown that testosterone and other androgens may have two different effects. Under oestrogen deprived conditions androgens after aromatase conversion may stimulate tumour growth via ER-α and this effect can be blocked by anti-oestrogens. In the presence of oestrogens, androgens inhibit the growth stimulatory effect of oestrogens and this antagonistic affect mediated via the androgen receptor can be blocked by anti-androgens [164]. AR is expressed in a significant subset of both ER-positive (89%) and ER-negative (49%) breast cancers [160] and its expression has been associated with a good prognosis in ER-negative/PR-negative breast tumours. Loss of AR expression, on the other hand, has been associated with poor prognosis in ER-negative/PR-negative/HER2-negative lymph node positive breast tumours [160].

Progesterone has been hypothesised based on *in vitro* studies and animal work to both decrease and increase breast cancer risk [165]. Epidemiological studies though have shown that progesterone in combination with oestrogens increases breast cancer risk [166]. Clinically, PR-A and PR-B are routinely assessed along with ER-α expression and overall it has been suggested that PR expression is positively associated with ER-α expression [160].

Regarding gonadotrophins there are not that many studies investigating their association with breast cancer. It has been shown though that gonadotrophins can act on breast cancer cells and accelerate conversion of DHEA into oestrogens stimulating the development of oestrogen-dependent tumour cells [167].

Exogenous Sex Steroid Hormones

OCs: OCs were first introduced in the 1960s and since then different formulations have been used. Most of them contain ethinyl oestradiol with new generation OCs containing lower concentration of oestrogen and new types of progestins. For a long time, it was hypothesised that OCs increase breast cancer risk but studies failed to show an association [168, 169]. A meta-analysis combining 54 epidemiologic studies provided the first convincing evidence that current OC use was associated with a 25% increase in breast cancer risk [170]. Studies looking into different doses and regimens of OCs to investigate whether they have a different effect on breast cancer showed that lower doses of newer regimens of OCs are associated with a lower breast cancer risk among young women [171].

HRT: Hormone replacement therapy was first introduced in the 1930s to manage menopause symptoms [172]. Early reports indicated that oestrogen alone was associated with an increased risk of endometrial cancer and this led to the addition of a progestagen. Increases in venous thromboembolic events were reported, but presumed beneficial effects for cardiovascular disease and osteoporosis and general well-being led to the continuation of its use [172, 173]. The first reports of an increased breast cancer risk came in 1976 from Hoover and colleagues [172]. Since then, several studies were carried out with two meta-analyses showing increased breast cancer risk among users of HRT [174, 175]. When the Collaborative group on breast cancer presented data showing an association between breast cancer and HRT, recommendation of hormonal therapy was taken more seriously [176] and the report published from the Women's Health Initiative Study caused HRT administration to drop quickly [177, 178]. A larger observational study in UK, the Million Women Study, confirmed the results from

previous studies showing that oestrogen in combination with progestin to be associated with higher breast cancer risk than oestrogen alone. The risks were also similar irrespective of formulation and sequential or continuous use of oestrogen or progestin [179].

Endogenous Sex Steroid Hormones

Reproductive Factors: Long duration of lactation and breastfeeding are associated with reduced risk of breast cancer and older age at first birth, decreased parity and late menopause are all well-established factors suggested to increase breast cancer risk [56, 180]. Breast cancer risk increases by almost 3% for each year delay to menopausal status (natural or surgical) [176] due to increased number of menstrual cycles [181]. Breastfeeding has a protective role, reducing breast cancer risk by 4% for every 12 months of breastfeeding [182]. There is 7% reduction in risk with each full term pregnancy, and overall women with children have 30% lower risk than nulliparous women [182]. Data has shown that reduction in breast cancer risk with childbirth and higher risk with later age at first full time birth may only be associated with ER-positive breast tumours [183]. Induced abortion and recognised spontaneous abortion were not associated with breast cancer [184].

Surrogates for lifetime sex steroid hormone exposure: Surrogates for long term sex steroid hormone exposure have also been suggested to be associated with breast cancer risk. Several anthropometric factors including height, weight, weight changes, body mass index (BMI), fat deposition and breast size [185] have been investigated in order to examine their association with breast cancer. Increased height has been associated with increased breast cancer risk

especially among postmenopausal women. Weight gain or BMI are factors that have been shown to have a different impact on risk depending on menopausal status. In premenopausal women increased weight decreases the risk but in postmenopausal women increased weight increases breast cancer risk [185, 186]. Weight gain during adult life increases postmenopausal breast cancer risk with the greater risk being observed when it occurs in the upper part of the body [185]. Fat adiposity which is measured by waist circumference and waist-hip ratio has been shown to be associated with the breast cancer risk in postmenopausal women but not in premenopausal. Weight-loss after menopause and physical activity has been shown to reduce breast cancer risk [180, 185]. Finally, breast size has also been investigated with conflicting results. The biological mechanisms behind the association of the above anthropometric measures and breast cancer risk include the nutritional lifestyle during childhood and adult periods, genetic predisposition, prenatal factors, IGF levels [185] and endogenous sex steroid hormones [187]. In a recent study these factors were shown to have an impact on tumour biology and pathology, confirming previously described associations between weight and increased risk of postmenopausal breast cancer and high BMI being associated with tumours expressing several markers corresponding with less aggressive cancers [188].

Another parameter could affect a woman's exposure to sex steroid hormones is diet. Meta-analysis of prospective studies investigating whether specific types of food consumption are associated with breast cancer risk showed a small increase or no associations [189, 190]. Phytoestrogens which are known to affect the hormone metabolism and bind to ER have been hypothesised to increase breast cancer risk but data from different studies are inconsistent [191]. The only well

established diet related risk factor is alcohol consumption which is known to increase bioavailable oestrogen and ethanol and may stimulate carcinogenesis by inhibiting DNA methylation [192]. Data from six prospective studies showed that alcohol intake is correlated with breast cancer incidence in women who drink alcohol regularly and reduction of consumption could lower the risk [193]. The relationship between smoking and breast cancer still remains controversial and for some studies it is considered to be a risk [194], whereas for others seems to have a protective effect lowering breast cancer risk [195] .

Circulating sex steroid hormones: Endogenous, exogenous sex steroid hormones and surrogates for long term sex steroid exposure have been shown to be associated with breast cancer as discussed above. Based on these observations it was hypothesised that circulating sex steroid levels may be a good measure of risk prediction for the total hormonal exposure that influences a woman's risk. In a study investigating the association of sex steroid hormones along with several epidemiological factors and breast cancer risk it was shown that oestradiol increased and SHBG decreased with increasing body mass index and the latter decreased with increasing waist-hip ratio. No associations were observed between sex hormones and age at menarche, parity, age at menopause, and previous use of oral contraceptives. Based on these observations it was suggested that obesity and perhaps waist-hip ratio may mediate their effects on breast cancer risk by changing circulating concentrations of sex hormones [196].

Table 2-2 summarises previous studies that have been carried out in order to relate endogenous circulating sex steroid hormonal levels with breast cancer risk

in postmenopausal women. A meta-analysis by Key *et al*, 2002 investigated the association of sex steroid hormones with postmenopausal breast cancer risk, combining 9 prospective studies, showed that women who have sex steroid hormone serum levels in the highest quintiles to have a two-fold increased risk of developing breast cancer [197]. One of the largest studies examining 663 cases and 1765 controls found total oestradiol, free oestradiol, oestrone, and oestrone sulphate to be associated with breast cancer risk [197]. Similar data was reported in EPIC study among 677 cases and 1309 controls [198]. Additionally, in the New York University Women's Health Study, two samples from one woman were analysed - one within 5 years of diagnosis and the second at least 5 years post diagnosis - to assess any changes in the hormone levels over time. The changes observed between the serial samples were not statistically significant different suggesting that circulating sex steroid hormones are a marker rather than a tumour related hormonal effect [199]. The same group also examined whether increased risk for DCIS is associated with high levels of sex steroid hormones but no significant trend was observed for any of the hormones examined [200]. In a more recent study it was shown that only oestrone and oestrone sulphate and not oestradiol and bioavailable oestradiol were associated with statistically significant increases in breast cancer risk [201].

There are conflicting data on endogenous levels of androgens and breast cancer risk [201]. In the largest case-control study high androgen levels were associated with higher breast cancer risk in postmenopausal women [197] which was supported by the EPIC study [198]. In general, it has been thought that the causal relationship of androgens in breast cancer is difficult to establish since increased aromatase activity in the setting of oestrogen depletion after menopause and

increased capacity to convert testosterone to oestradiol may be both the major factors. After adjustment for oestradiol the association of the androgens with breast cancer risk remained indicating that androgens have an effect independent of the oestrogens [197, 198, 202].

Interestingly, despite the association of exogenous progesterone and breast cancer risk, the association of endogenous progesterone and breast cancer risk is not clear. There is only one large study among postmenopausal women by Missmer *et al*, which has not shown any association of the circulating progesterone and breast cancer risk [202].

The last few years there has been an interest in assessing the association of sex steroid hormones and breast cancer risk by oestrogen and progesterone receptor status of the tumour. The first report was that by Helzlsouer *et al*, demonstrating that in postmenopausal women the association of endogenous oestrogens with breast cancer risk was independent of the ER status of the tumour [203]. Almost 10 years later, another study reported that circulating levels of sex steroid hormones were most strongly associated with risk of ER-positive/PR-positive breast tumours [202]. In a recent larger study positive association was observed for oestradiol and testosterone for ER-positive/PR-positive tumours and weak and no association for ER-positive/PR-negative and ER-negative/PR-negative tumours [202]. A recent study confirmed these results providing further evidence that the developing tumours are mainly oestrogen receptor positive and showed that although HER2-positive and HER2-negative breast cancers were both associated with high total testosterone, they showed opposing associations with oestrogen [204]. More recently, a study by Baglietto *et al*, showed conflicting results

reporting that the associations of endogenous hormones with postmenopausal breast cancer risk are independent of tumour grade, and hormone receptor status but that might increase in strength with age [205].

Studies have also looked into whether there is an association between sex steroid hormones and breast cancer in women at varying levels of breast cancer risk. No association was observed between reduced risk in tamoxifen treated women in the high risk population of the National Surgical Adjuvant Breast and Bowel Project Cancer Prevention trial with both androgens and oestrogens levels [206]. In the Nurses's Health Study cohort women with high levels of endogenous oestrogens and testosterone were at increased breast cancer risk regardless of predicted risk or family history of breast. Therefore, it was suggest that sex steroid hormones are predictive of risk irrespective whether a woman has an increased predicted breast cancer risk [207].

All the above described studies have included only women who were not using HRT treatment. In order to investigate associations between sex steroid hormones and breast cancer risk among women using hormonal treatment and those who do not a prospective study was carried out. The data suggested that higher sex steroid hormone levels are associated with breast cancer among the hormonal users [208].

Studies have also been carried to investigate premenopausal breast cancer risk in association with sex steroid hormones. That relationship still is not clear as these studies are difficult to be carried out due to large intra-individual variation related to menstrual cycle. The two larger studies carried out since now are from

European Prospective Investigation into Cancer and Nutrition (EPIC) study demonstrating only androgens to be associated with premenopausal breast cancer risk [209] and the Nurse's Health Study reporting both oestrogens and androgens to be associated with premenopausal breast cancer risk [210]. For the purposes of this thesis, the literature review presented is focused on postmenopausal women.

2.2.7 Gonadotrophins and breast cancer

Oestrogen synthesis is under the control of LH and FSH. LH through its receptor stimulates the production of ovarian androgens and FSH the aromatisation of androgens to oestrogens. Therefore, it is of great interest to investigate whether LH and FSH are associated with breast cancer risk. To date, studies measuring the actual LH and FSH levels have failed to demonstrate an association with breast cancer risk [211, 212] in contrast to mouse model work demonstrating that over expression of LH is responsible for the development of spontaneous mammary tumours that lack PR expression [213]. High levels of FSH and LH have been shown to stimulate both normal and malignant human ovarian surface epithelial cell growth [214] and FSH has been reported to be associated with ovarian cancer risk [215]. Further studies are needed to better understand the role of FSH and LH in breast cancer. It would be of great interest to investigate the association of gonadotrophins with HRT use since previous studies have suggested that HRT stimulates the growth of only the clinically significant breast cancers and is known to increase breast density (risk factor) reducing the sensitivity on the other hand of mammography [216].

Table 2-2: Summary of studies investigating association between endogenous sex steroid hormones and postmenopausal breast cancer risk.

Author and Year	Study Design and No of patients	Assays and Hormones Analysed	General Findings
Helzlsouer et al, 1994 [203]	Nested Case-Control Study 130 cases and 260 controls	Free E2 assay: ultra filtration method, Total E2 and E1: RIA, FSH: radiometric assay	Association of endogenous oestrogens with breast cancer risk is independent of the ER status of the tumour.
Toniolo et al, 1995 [212]	Nested Case-Control study 130 cases and 260 controls (Mean age of cases 58.9 yrs) Invasive breast cancer, no comment on histology	Free E2: ultrafiltration method, total E2 and E1: RIA, FSH: standard radiometric assay	First confirmation in a large prospective epidemiologic study of a link between circulating oestrogens and breast cancer risk. Higher levels of oestrone, total oestradiol, and free oestradiol, and a lower percent of oestradiol bound to SHBG for women who developed breast cancer than women who remained free of cancer.
Key et al, 1996 [217]	Case Control Study 69 developed breast cancer after joining the study (Mean age of cases 58.3 yrs) No comment on histology	RIA assays: oestrone, oestradiol, oestriol-urine samples	High levels of endogenous oestrogens in postmenopausal women are associated with increased breast cancer risk but that the relationship of oestrogens in premenopausal women with risk unclear.
Dorgan et al, 1996 [218]	Prospective Nested Case Control Study 72 postmenopausal women and 144 controls (matched: on age and on date and time of day) (Median age cases:61 yrs and controls: 62 yrs)	E2, testosterone, androstenedione, DHEAS, E1S: RIA after extraction, SHBG: immunoradiometric assay	Further evidence in support of the already established association between elevated oestrogen levels and breast cancer. New evidence that high serum testosterone levels precede breast cancer occurrence.
Berrino et al, 1996 [219]	Case Control Study 24 cases and 88 controls (Mean age cases: 59.4 yrs and controls: 54.9yrs)	E1 and Total testosterone: a non-extraction RIA, free testosterone: coated-tube RIA, SHBG: immunoradiometric.	Further evidence in support of the already established association between elevated oestrogen levels and breast cancer. Evidence that high serum testosterone levels precede breast cancer occurrence.
Dorgan et al, 1997 [220]	Nested Case Control Study 72 cases and 144 controls (Median age cases:61 yrs and controls: 62 yrs) No comment on histology	RIA: E2, testosterone, E1, androstenedione and DHEAS	Risk of breast cancer was positively and significantly associated with serum levels of oestrogens and androgens. The results lend considerable support to the hypothesis that serum concentrations of oestrogens and androgens are related to the subsequent diagnosis of breast cancer in postmenopausal women.

DHEAS=dehydroepiandrosterone sulphate; E2=oestradiol; E1=oestrone; E1S=oestrone sulphate; FSH=follicle stimulating hormone; RIA=radioimmunoassay; SHBG=sex hormone-binding globulin

Author and Year	Study Design and No of patients	Assays and Hormones Analysed	General Findings
Dorgan et al, 1997 [221]	Prospective Nested Case Control Study 72 cases and 144 controls (Median age cases:61 yrs and controls: 62 yrs) No comment on histology	DHEA and DHEAS: RIA with extraction	Increasing risk of breast cancer for increasing concentrations of DHEA. The relationship of DHEAS to breast cancer was less consistent, but women whose serum DHEAS concentration was in the highest quartile also exhibited a significantly elevated risk ratio.
Thomas et al, 1997 [222]	Case Control Study 61 cases and 179 controls No comment on histology	RIA	High concentration of E2 associated with breast cancer. Adjustment with testosterone and SHBG concentrations did not change the odds ratio for E2. Testosterone and SHBG concentrations were associated with breast cancer risk but the associations were not statistically significant after adjusting for E2. Evidence that serum oestradiol concentrations in postmenopausal women may have a substantial effect on breast cancer risk.
Zeleniuch-Jacquotte et al, 1997 [223]	Nested Case Control Study within the New York University Women's Health Study 130 cases and 260 controls (Mean age cases: 59.2 yrs and controls: 59.1 yrs)	Total testosterone: solid-phase RIA: DHEAS directly in diluted serum, total E2: standard RIA and % E2 bound to SHBG and % E2 free: concanavalin A-sepharose binding and an ultra filtration method respectively	Testosterone associated with breast cancer risk but after adjustment with E2 bound to SHBG and total E2 no longer significant. Breast cancer risk remained associated with total E2 levels and negatively associated with E2 bound to SHBG after adjustment for serum testosterone levels. No evidence was found of an association between DHEAS and risk of breast cancer in postmenopausal women.
Hankinson et al, 1998 [224]	Nested Case Control Study within the Nurses' Health Study 156 cases and 312 control (matched with respect to age, menopausal status, month and time of day of blood collection, and fasting status at the time of blood collection)	RIA	Strong evidence for a causal relationship between postmenopausal oestrogen levels and the risk of breast cancer.

DHEA=dehydroepiandrosterone; DHEAS=dehydroepiandrosterone sulphate; E2=oestradiol; RIA=radioimmunoassay; SHBG=sex hormone-binding globulin

Author and Year	Study Design and No of patients	Assays and Hormones Analysed	General Findings
Cauley et al, 1999 [225]	Prospective Case Cohort Study. 97 cases and 244 controls; not receiving oestrogen therapy (Mean age cases:70.9 yrs and controls:71.8 yrs)	Total E2, E1, E1S, androstenedione, DHEA: RIA after extraction and chromatography, free E2: equilibrium dialysis and calculated by the % of dialyzable oestradiol, total E2, % of non-SHBG-bound E2 or bioavailable E2: monium sulphate precipitation of SHBG-bound steroids, total testosterone: RIA with chromatographic purification, free testosterone: equilibrium dialysis, SHBG: direct RIA	Oestradiol and testosterone levels may play important roles in the development of breast cancer. A single measurement of bioavailable oestradiol and free testosterone may be used to estimate a woman's risk for breast cancer. Women identified as being at high risk for breast cancer as determined by these hormone levels may benefit from antioestrogen treatment for primary prevention.
Kabuto et al, 2000 [226]	Case Control Study from the Life Span Study, Japan 72 cases and 150 control (matched on age, date of blood collection, exposure, radiation dose) (Mean age cases: 60.7 yrs)	E2, prolactin, SHBG and progesterone: RIA, levels of bioavailable E2: calculated, DHEAS: RIA for 11-deoxy-17 ketosteroid	Further prospective support for the hypothesis that a high level of biologically available E2 is a risk factor for the subsequent development of breast cancer.
Key et al, 2002 [197]	Meta-analysis of 9 studies 663 cases and 1765 controls	Different assays applied within the different studies; oestradiol, free oestradiol, bioavailable oestradiol, oestrone, oestrone sulphate, testosterone, andostenedione, DHEA,DHEAS, SHBG	All hormones statistically significant associated with an increased breast cancer risk.

DHEA= dehydroepiandrosterone; DHEAS=dehydroepiandrosterone-sulphate; E2=oestradiol; E1=oestrone; E1S=oestrone sulphate; RIA=radioimmunoassay;

SHBG=sex hormone-binding globulin

Author and Year	Study Design and No of patients	Assays and Hormones Analysed	General Findings
Manjer et al, 2003 [211]	Two populations based prospective cohort studies in Sweden. Blood samples were collected in about 65,000 women/Follow-up yielded 173 postmenopausal breast cancer cases who had not been exposed to HRT (Mean age cases: 61.6 yrs and controls: 60.5 yrs)	Testosterone, androstenedione: competitive RIA, DHEAS, E1, E2: direct RIA, FSH, prolactin: sandwich magnetic separation assay, SHBG: immuno fluorometry	High levels of E1, E2, testosterone, and possibly androstenedione and DHEAs, in postmenopausal women are associated with a high risk of subsequent breast cancer.
Onland-Moret et al, 2003 [227]	A nested case-cohort study was conducted within a large cohort (the DOM cohort) in the Netherlands (n=9,349) Women using hormones were excluded leaving 364 breast cancer cases and 382 women in the cohort for the analyses (Mean age cases: 61.6 yrs and controls 60.5 yrs)	E1, E2, testosterone and 5a-androstane-3a, and 17b-diol (3aD): RIA after extraction and chromatography	Women with higher excretion levels of both oestrogens and androgens have an increased risk of breast cancer.
Lamar et al, 2003 [228]	A cross-sectional study 133 women (Mean age cases: 61 yrs and controls: 62 yrs)	E2, testosterone, androstenedione, DHEAS, E1S: RIA after extraction and chromatography, SHBG: immunoradiometnic assay, % of unbound and albumin-bound E2: centrifugal ultra filtration	Higher oestrogens and possibly testosterone mediate the increased breast cancer risk associated with obesity. Higher testosterone levels could potentially contribute to the increased risk of breast cancer among nulliparous postmenopausal women. The results did not support a role for changes in serum oestrogen, androgen and SHBG levels in explaining the age-related increase in breast cancer incidence.

DHEAS=dehydroepiandrosterone-sulphate; E2=oestradiol; E1=oestrone; E1S=oestrone sulphate; FSH=follicle stimulating hormone; RIA=radioimmunoassay; SHBG=sex hormone-binding globulin

Author and Year	Study Design and No of patients	Assays and Hormones Analysed	General Findings
Missmer et al, 2004 [202]	Nested Case Control Study 264 invasive, 41 in situ,153 ER+/PR+, 39 ER-/PR- cases and 643 controls Mean age of cases: 62 yrs	E1, E2, testosterone, SHBG, DHEAS, progesterone: RIA and free and percent free E2: calculated by the law of mass action	Circulating levels of sex hormones strongly associated with risk of ER+/PR+ breast tumours.
Zeleniuch-Jacquotte et al, 2004 [199]	Nested Case Control Study 297 cases and 563 controls Mean age of cases: 60 yrs	E1, E2, androstenedione, FSH, testosterone and DHEAS: direct RIA, SHBG: direct 'sandwich' immunoradiometric assay	Associations of circulating oestrogen with breast cancer risk are more likely due to an effect of circulating hormones on the development of cancer than to elevations induced by the tumour. Contribution of androgens to risk is largely through their role as substrates for oestrogen production.
Tworoger et al, 2005 [208]	Nested Case Control Study 446 cases and 459 controls Mean age of cases: 59.9 yrs and controls: 59.8 yrs	E2 and testosterone: RIA after extraction and chromatography, SHBG: immunoassay, free E2 and free testosterone: calculated by the law of mass action	Although women using hormonal treatment have a different hormonal profile than those not using hormonal treatment plasma sex hormone concentrations are associated with breast cancer among those who were treated with hormones.
Kaaks et al, 2005 [198]	Nested Case Control Study 677 cases and 1309 controls Mean age of cases: 60.4 yrs and controls 60.3 yrs	Testosterone and DHEAS: RIA, androstenedione, E1 and E2: RIA with a double-antibody system for the separation of free and bound antigen	Elevated serum oestrogens and androgens associated with increased breast cancer risk. Since DHEAS and androstenedione are largely of adrenal origin in postmenopausal women, the results indicated that elevated adrenal androgen synthesis is a risk factor for breast cancer. Caution against the use of DHEA(S) or other androgens for postmenopausal androgen replacement therapy.
Zeleniuch-Jacquotte et al, 2005 [200]	Nested Case Control Study 69 DCIS and 134 matched controls Mean age of cases: 58 yrs and controls 58 yrs	E2, E1 and androstenedione: direct double-antibody RIA, testosterone and DHEAS: direct RIA, SHBG: direct sandwich immunoradiometric assay, FSH: immunometric assay	No statistically significant trend of increasing risk with increasing level of any hormone was observed.

Author and Year	Study Design and No of patients	Assays and Hormones Analysed	General Findings
Adly L et al, 2006 [201]	Case Control Study – samples taken at diagnosis 179 cases (invasive breast cancer and 152 controls (benign conditions of breast) Mean age of cases: 67.1 yrs and controls	E2, E1, E1S testosterone, DHEA and androstenedione: RIA after extraction and chromatography	Higher serum concentrations of oestrogens were associated with increased breast cancer risk.
Eliassen et al, 2006 [207]	A prospective nested case-control study within the Nurses' Health Study 418 cases 817 age matched controls Low/Moderate/High Risk of cases based on Gail Rosner and Colditz models	RIA following extraction and celite chromatography	Higher levels of endogenous oestrogens and testosterone are associated with increased breast cancer risk regardless of predicted risk or family history of breast cancer.
Beattie et al, 2006 [206]	Case-Cohort Design 135 cases and 275 controls Women had enrolled in the National Surgical Adjuvant Breast and Bowel Project Cancer Prevention Trial and who had been treated with tamoxifen or placebo for 69 months	RIA: oestradiol, testosterone, SHBG	Reduced risk of invasive breast cancer in tamoxifen-treated women compared with placebo treated women was not associated with sex steroid hormone levels. The data did not support the use of endogenous sex hormone levels to identify women who are at particularly high risk of breast cancer and who are most likely to benefit from chemoprevention with tamoxifen.
Sieri et al, 2009 [204]	Case-Control Study 165 who developed breast cancer after being followed up for 13.5 years ER+, ER-, PR+, PR- Mean age cases: 58.02 yrs and controls 58.10 yrs	RIA: testosterone and oestradiol and immunoassay: SHBG	High levels of circulating testosterone increase the risk of postmenopausal women to develop breast cancer. The cancer that they developed was mainly oestrogen receptor positive.
Baglietto et al, 2010 [205]	Case Study 197 postmenopausal women with breast cancer and 857 random chosen women	Testosterone and E2: electrochemiluminescence immunoassay, E1S and androstenedione: RIA, DHEAS: competitive immunoassay, SHBG: immunometric assay	Associations of endogenous hormones with postmenopausal breast cancer risk are independent of tumor grade and hormone receptor status and may increase with age.

DHEAS=dehydroepiandrosterone-sulphate; E2=oestradiol; E1=oestrone; E1S=oestrone sulphate; FSH=follicular stimulating hormone; RIA=radioimmunoassay;

SHBG=sex hormone-binding globulin

2.3 Epigenetics

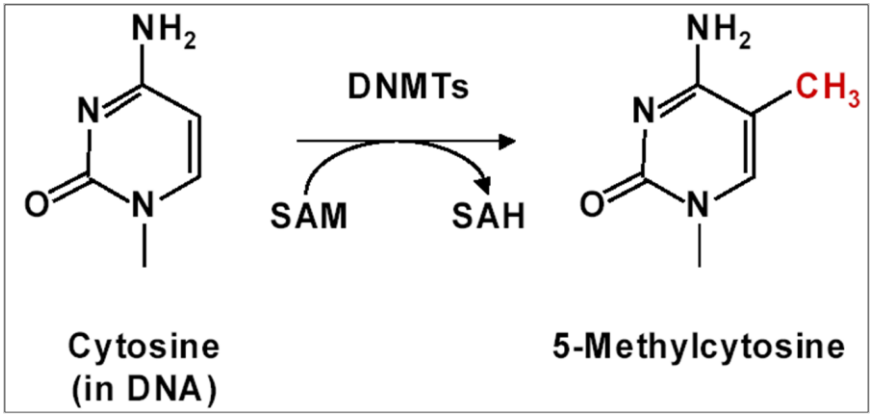
Epigenetics have been defined as modifications of DNA or associated factors with information content other than the DNA sequence itself that are maintained during cell division, mitosis and/or meiosis [229]. The Greek 'epi'- prefix of the word 'epigenetics' implies features that are 'on top of' or 'in addition to' genetics. Therefore, the term epigenetics encompasses events that influence gene function, but it is on top of or in addition to the traditional molecular basis for inheritance. The term was first introduced in 1940s describing the interaction between genes and environment in the development of specific phenotypical traits, which cannot be explained by genetic principles. There are four main, inter-related types of epigenetic inheritance which are all linked together acting in a synergistic way: DNA methylation, histone modifications, nucleosome positioning and non-coding RNAs, specifically microRNA expression. Epigenetic modifications are known to be an early event in carcinogenesis and to precede major genetic changes leading to cancer [115]. Several reports have demonstrated an association between DNA methylation changes and breast cancer (discussed in section 2.4.3). Women with gene specific DNA methylation changes are at an increased risk to develop the disease with odds ratio (OR) ranging from 1.4 to 5.28 [163, 230].

2.3.1 DNA methylation

DNA methylation refers specifically to the covalent addition of a methyl group from the methyl group S-adenosyl-L-methionine (SAM) to the carbon-5 position of the cytosine ring to form the called fifth base, 5-methylcytosine [231]. The reaction is catalysed by a family of enzymes which are transferring the methyl group from the donor molecule, SAM, to the cytosine ring known as DNMTs [232, 233] (Figure 2-

3). Several distinct physiologically active members have been cloned and characterised including DNMT1 which is responsible for maintaining methylation after DNA replication and DNMT3a and DNMT3b which are responsible for *de novo* methylation during early embryogenesis. Studies have shown though that DNMT1 is not sufficient in maintaining methylation with *de novo* activities of DNMT3a and b being necessary for the establishment of methylation patterns in the genome [234].

Figure 2-3: DNA methylation reaction catalysed by DNA methyltransferases.



But how DNMTs are targeted in particular sites within the genome causing DNA methylation? One of the suggested mechanisms is by the recognition of specific chromatin structures. Chromatin is known to consist of 146 base pairs of DNA wrapped around an octamer of four core histone proteins: H3, H4, H2A and H2B. It is found in two states either in an active euchromatic state or an inactive heterochromatic state. Euchromatin is defined by di- and trimethylation of lysine 4 on histone H3 and acetylation of histones H3 and H4. Heterochromatin is characterised by either trimethylation of lysine 27 and 9 on histone H3 or methylation of lysine 20 on histone H4 [235]. These modifications are known to be regulated by enzymes that add and remove covalent modifications to histone

proteins. Based on the modifications the proteins are divided into: histone acetylases (HATs), methyltransferases (HMTs), deacetylases (HDACs), demethylases (HDMs). Studies have suggested that DNA methylation occurs at heterochromatic regions, these histone modifications either individually or in combination make possible targets for the DNMTs. DNMTs have also been shown to interact with HMTs such as G9a, protein arginine methyltransferase 5 and SUV39 [115]. SUV39 is responsible for methylation of lysine 9 on histone 3 and the enhancer of zeste homologue 2 (EZH2) which is one of the proteins contained in the Polycomb Repressor Complex (PRC) 2 also containing the embryonic ectoderm development (EED) and suppressor of zeste homologue 12 (SUZ12), which catalyses the methylation of lysine 23 on histone H3. PRC1 and PRC2 complexes are known to regulate gene expression of embryonic stem (ES) cells which contain coexisting active trimethylated lysine 4 on histone H3 and repressive trimethylated lysine 27 on histone H3 marks at the promoters of genes that are important in developmental processes [236]. Recently it has been suggested that the occurrence of methylation and dimethylation of lysine 9 on histone H3 and trimethylation of lysine 27 on histone H3 within the same region of the genome can serve as a signal to recruit DNMTs [237]. DNMTs are also known to interact with heterochromatin protein I, a protein that is known to specifically bind to methylated lysine 9 on histone 3 [115, 236]. Other suggestions include the recruitment of DNMTs by repressors and RNAi with conflicting data being presented and further studies needed to show that this mechanism is important in mammals [236].

Methylation occurs predominantly in cytosines located 5′ of guanines and known as CpG dinucleotides, where p refers to phosphate link between the two

nucleosides. CpGs are unequally distributed and they are greatly underrepresented in the mammalian genome through evolutionary loss of 5-methylcytosines through deamination to thymine [238]. However, clusters of CpGs known as CpG islands (CGIs) are present in 1–2% of the genome (approximately 30,000 CGIs have been suggested to be present in the genome) and their length ranges from 200 bp to 2 kilo bases. CGIs are frequently contained within and around the promoter regions, in the first and second exons and the first intron, of the mammalian gene and it has been estimated that around 40% of all genes contain a CGI [239]. The fact that CGIs are localised in the promoter region of a gene makes them critical in gene regulation, usually with an inverse relationship between the degree of methylation of a regulatory CGI and the extent of gene transcription [240].

The regulation of gene expression by DNA methylation modifications is based on the suggestion that DNA methylation is able to physically prevent the binding of transcription factors to their binding sites in the promoter of the genes, therefore inhibiting the transcription process. Another mechanism is based on the theory that DNA methylation can prevent transcription by interfering with the propagation of active chromatin marks. Studies provided evidence that methylated DNA is able to recruit a family of proteins known as methyl-CpG binding proteins (MBDs) consisting of five well-characterised members: MeCP2, MBD1, MBD2 MBD3 and MBD4. These proteins are important mediators between DNA methylation and histone modifier genes establishing a transcriptionally inactive chromatin through their association with protein complexes that involve the action of HDCA1 and HDCA2 and chromatin remodelling proteins such as sin3a and mi-2. This protein association is responsible for deacetylation of the histones that leads to a tighter

binding between the positively charged lysine residues of histones and the negatively charged phosphodeoxyribose backbone of the DNA reducing accessibility of DNA for transcription factors [115, 236]. Even though there are a lot of studies trying to better understand the molecular interplay between these epigenetic modifications, still the mechanisms which underlie the link between DNA methylation and histone modifications remain under intense scrutiny. The hierarchy and chronology of DNA methylation, histone modifications and altered gene transcription are yet to be established.

2.3.2 DNA methylation detection methods

Detection of DNA methylation changes is based on the ability to differentiate between cytosine and 5-methylcytosine in the DNA sequence. Nowadays, there are a variety of methods which can be used to obtain DNA methylation data. These methylation techniques can be categorised according to the following DNA treatments: 1) methylation sensitive restriction digestion, 2) immunoprecipitation and 3) sodium bisulphite modification (BM).

Methylation sensitive restriction digestion analysis using specific restriction enzymes is a technique that has been used for many years to confirm the methylation status of CpG dinucleotides [241, 242]. Due to increased interest in developing methods which can examine genome-wide epigenetic alterations restriction landmark genomic scanning was introduced in 1991 [243] allowing single base resolution via sequencing. More recent developments such as differential methylation hybridisation via CpG-island microarrays provide an attractive alternative [244]. The drawback of these techniques is that they are significantly labour intensive and require high concentrations of DNA.

Immunoprecipitation is another approach. Main advantage of the technique is the lack of requirement for restriction digestion that reduces sequence bias. It involves two approaches: 1) Methylated-CpG Island Recovery Assay (MIRA) which uses antibodies against the MBD family of proteins that preferentially binds to methylated DNA. MIRA has been developed in conjunction with CpG island arrays, and very recently, was used to demonstrate increased methylation of homeobox genes in breast cancer [245, 246], 2) MeDIP which involves the use of a monoclonal antibody directly against methylated cytosines [247].

The most widely used techniques and those that have been at the forefront of DNA methylation analysis are those that involve the use of chemically treated DNA with sodium bisulphite. Sodium BM relies on the differential deamination of cytosine to uracil without affecting the 5-methylcytosine content [248]. The conversion produces differences in the DNA sequence which are dependant on the original methylation status of the genome. These differences can be used to design PCR primers which will either amplify a region depending on its methylation status or amplify a pool of unmethylated and methylated products. The most well known method is methylation specific PCR (MSP) which was introduced in 1996 [249]. Several techniques since then have been developed including MethyLight [250], combined bisulphite restriction analysis (COBRA) [251] and pyrosequencing [252].

MethyLight has an increased level of sensitivity as a result of the incorporation of a probe with the primers; however, this can complicate assay design. Quantification of methylation by MethyLight is represented by the ‘percentage of fully methylated reference or ‘‘PMR’’ which compares the fluorescence intensities

of the target gene in the sample with those of a theoretically fully methylated reference DNA [250]. COBRA and pyrosequencing provide quantitative information using PCR-primers that do not cover any potentially methylated CpG sites. COBRA relies on a methylation sensitive restriction enzyme digest to provide quantitative assessment of the methylation status of individual CpG sites [251]. Pyrosequencing which is an improved method of bisulphite genomic sequencing provides assessment at single CpG dinucleotide level [252]. Recently, pyrosequencing was re-introduced as the 'Next-generation' sequencing involving PCR amplification of target DNA and use of fluorophores, a method during which incorporation of each nucleotide is accompanied by an enzymatically driven emission of light. This process was recently used to perform massively parallel bisulphite sequencing from serum and breast tissue [253].

Until now, the majority of studies have relied on a candidate gene approach allowing the analysis of a limited number of genes. Very recently, epigenome wide analyses came into the scene with Illumina introducing their universal bead array technology in the form of the Goldengate and Human Methylation 27 (Methyl 27K) platforms [254, 255]. Both of them generate quantitative data expressed as beta (β) which are continuous variables between 0 and 1, representing the ratio of the intensity of the methylated bead type to the combined locus intensity. The beadchip technology allows detection of methylation levels down to as little as 2.5% [255] and the Illumina Methyl 27K platform has a capacity for the simultaneous analysis of approximately 14,000 genes. As described there are limitations in any of the methods available and no single technique can be considered better than the other one. Validation of the data is important either by

using a second set of samples. The application of these tools in clinical research is critical in breast cancer as it will allow not only to identify novel methylation targets but it will also enable the identification of patients that could eventually benefit from treatment.

2.3.3 DNA methylation during normal development and disease

In humans DNA methylation patterns are established during defined phases in embryonic development. After fertilisation dramatic waves of methylation changes occur. Gamete methylation patterns are erased by a genome-wide demethylation event at around the eight-cell stage of blastocyst formation [256]. During implantation, DNA methylation patterns are re-established via *de novo* methylation and are maintained through subsequent cell divisions [257]. During adulthood, the primary role of DNA methylation is the maintenance of transcriptionally silent repetitive DNA elements which are scattered all over the human genome preventing chromosomal instability [258]. In contrast, most CGIs are unmethylated under normal circumstances in normal tissue [259], with the exception of those associated with imprinted genes with promoter methylation of either the paternal or maternal allele [260] and genes subjected to X chromosome inactivation in females [247]. There are also studies showing methylated non-imprinted autosomal CGIs in normal cells playing an important role in the establishment and control of cell type specific gene expression e.g. Homeobox A5 *(HOXA5*) [261].

Disruption of these pre-set patterns of DNA methylation during adult life have been linked to aging and disease. Several congenital malignancies such as immunodeficiency, centromeric region instability, facial anomalies syndrome which have a mutation in DNMT3B enzyme have been shown to be associated

with hypomethylation [262]. Methylation changes have also been linked to Beckwith-Wiedemann and Prader-Willi syndromes which are imprinting disorders [263]. Furthermore, dysregulation of developmental programming by maternal and/or environmental factors is thought to induce abnormal DNA methylation of specific genes and thence their faulty expression, leading to disease [264]. Such observations have been made by studying epigenetic differences between monozygotic twins. In early age monozygotic twins it is not possible to distinguish any epigenetic differences but as they grow older several differences in their epigenome are seen suggesting that the influence of the environment is an important parameter that needs to be taken into account [265]. Age dependent methylation alterations are also observed in normal tissues [266] but the most significant and frequently studied changes are those detected in many cancer types including breast cancer.

2.3.4 DNA methylation and cancer

During carcinogenesis normal cells undergo an extensive epigenetic transformation. The cancer epigenome is characterised by global changes in DNA methylation and histone modification patterns and altered expression of chromatin modifying enzymes. Segregation of the epigenome into unmethylated and methylated regions is responsible for the formation of a rigid repressive chromatin which leads to reduced cellular plasticity. These changes result in dysregulation of gene expression profiles and along with genetic alterations play an important role in cancer initiation and progression of cancer. When gene expression is altered due to DNA methylation, it is usually characterised as due to hypomethylation or hypermethylation [115].

It was first shown that the genome of cancer cells is hypomethylated in comparison to normal tissue [267]. The genome-wide hypomethylation observed in cancer is mostly due to loss of methylation from repetitive elements in the genome resulting in genomic instability by promoting chromosomal rearrangements [268]. It is also responsible for the activation of oncogenes (growth-promoting genes) such as *c-myc* (C-myelocytic leukaemia) and loss of imprinting in colorectal cancer [269].

The most well characterised epigenetic modification though during carcinogenesis is the *de novo* methylation of CGIs around the promoter region of genes correlating with transcriptional repression. CGI methylation and subsequent transcriptional silencing occurs at least as often as genetic alterations in tumour suppressor genes in cancer [270]. Various tumour suppressor genes have been identified to undergo tumour-specific silencing by hypermethylation. These genes are involved in cell cycle regulation, apoptosis, transformation, signal transduction and adhesion, angiogenesis and metastasis [85]. Moreover, indirect silence of genes by silencing transcription factors and DNA repair genes has also been shown.

Despite the fact that we know a great deal regarding these hypo- and hyper-methylation changes in cancer still the events that lead to their initiation and the mechanism by which CGIs in normal cells are protected against methylation but lose this protective barrier in cancer and become hypermethylated are not fully understood. Epigenetic alterations have been suggested to be initiating events in the expansion of cells in preneoplastic lesions but the influences of these alterations as initiation events have been difficult to study. It is known that

methylation of specific genes alongside coordinated genetic hits potentially drive the development of a cancer, with multiple epigenetic hits being shown to be potential early events in precancerous cells prior to genetic alterations predisposing cancer cells to further mutations and increasing the likelihood of tumour progression [271]. Further to this, methylation in premalignant breast and colorectal tissue has been suggested to represent a field defect, perpetuating further neoplastic changes [272, 273]. There is also the recent suggestion that epimutations of stem cells may be the initiating progenitor event in tumourigenesis [264]. In addition, as epigenetic modifications are mitotically heritable they provide a growth advantage to rapidly growing cancer cells that result in their proliferation [115]. In addition, it has also been shown that epigenetic modifications are affected by age [30, 274] environment [275], chronic inflammation [276] and endocrine exposure [163].

Tumour-specific CpG island methylation has been suggested to occur through a sequence specific instructive mechanism during which DNMTs are targeted at specific genes through an association with oncogenic transcription factors [115]. Additionly, it has been shown that *de novo* methylation may start in exonic CGIs and subsequently spread into the promoter region of genes [277]. Alterations in methylation have been believed to locally silence discrete genes during carcinogenesis but recent work has challenged this concept by showing that long range epigenetic silencing may exist hypermethylating neighbouring genes and causing global gene silencing through chromatin remodelling activities [278].

Recent evidence has also shown that genes which are methylated in cancers may be vulnerable to aberrant DNA hypermethylation and epigenetic silencing

during tumour initiation and progression because of alterations in chromatin structure in stem or progenitor cells, including dimethylated and trimethylated lysine 9 on histone H3 [279, 280]. This finding supports the cancer stem cell hypothesis which is based on the observation that tumourigenic tissue contains a heterogeneous population of cells that are characterised by tumourigenic properties. As epigenetic modifications are key for the maintenance of stem cell identity it has been hypothesised that their disruption could give rise to a high risk aberrant progenitor cell population which is capable of undergoing transformations leading to the subsequent production of mutations. This phenomenon can lead to an overall increase in the number of progenitor cells and an increase in their ability to keep their stem cell state, forming a high risk population which can finally become neoplastic through additional genetic mutations [115].

Based on these suggestions a new model of carcinogenesis has been suggested. The predisposition of stem cell PRC2, which contains EZH2, EED and SUZ12 as mentioned before, targets to cancer-specific DNA hypermethylation suggesting a 'crosstalk' between PRC2 and *de novo* DNA methyltransferases in precursor cancer cells with a PRC2 target gene distribution similar to that of stem cells. This 'crosstalk' may be initiated and/or facilitated by various environmental exposures, transgenerational inheritance, endocrine exposure, inflammation and by age. A stem cell whose potential to differentiate has been irreversibly blocked by CpG methylation would then be predisposed to carcinogenesis via the acquisition of further genetic events, such as mutations and deletions (Figure 2-4) [280]. Better understanding of how specific genomic regions are targeted for DNA hypermethylation and how these DNA modifications are initiated in cancer will

potentially lead to therapeutic strategies and identification of biomarkers for early detection

Figure 2-4: Fixation of a stem cell signature by means of DNA methylation as prerequisite for carcinogenesis.

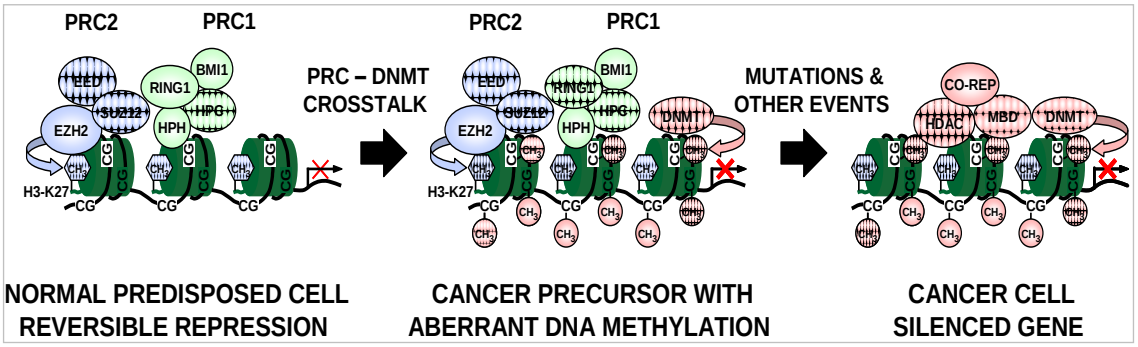


Figure from Widschwendter et al [280].

2.3.5 DNA Methylation biomarkers and breast cancer

Hypermethylation or hypomethylation of CGIs is a potentially attractive marker for detecting the neoplasm and detection of these changes have been proposed as a potential early diagnostic tool in cancer. Beside the presence of epigenetic alternations in the tumour tissue, DNA methylation changes in cancers can frequently be assayed in various sources of body fluids, serum and plasma, and may serve as a potential target to early detect cancer or to detect minimal residual disease after primary treatment has been completed. The precise mechanism by which DNA is released into the bloodstream still remains uncertain but evidence is accumulating that in areas of high cell turnover and cell lysis, DNA from necrotic and apoptotic cells can be transported from the intercellular space via lymph vessels into the blood stream [86]. It has already been shown by several groups that CGI methylation can be detected in plasma with the same characteristic changes as those found in the corresponding tumour and

cancerous tissues are methylated in contrast to CGIs in normal tissue making DNA methylation signature a promising biomarker [10]. Amongst the methylated genes are tumour suppressor genes such as *p16,* damage response genes such as *BRCA1*, mismatch repair genes e.g. *hMLH1* and *HMSH2*, steroid receptor gene family members such as *ER*, *PR* and retinoic acid, cell adhesion and cell surface molecules and DNMT inhibitors [85]. Several reviews have summarised breast cancer biomarkers [85, 281, 282] with a recent review specifically focusing on the presentation of markers discovered by examining breast tumour tissue [283]. Another review summarised studies that have revealed various genes to be either hypo- or hyper- methylated in breast cancer [284]. Recent diagnostic and risk prediction DNA methylation breast cancer markers are shown in Table 2-3 [116]. Markers were analysed in a variety of tissue sources including, tumour tissue, serum/plasma, peripheral blood cell DNA, nipple aspirate/duct fluid and fine needle aspirate washings.

DNA methylation analysis for the early detection of breast cancer was pioneered by Evron *et al*, 2001 by comparing methylation of cyclin D, *RAR-β* and *TWIST* promoters using cells extracted from ductal lavage fluid. In this study apart from detecting methylation changes in women with DCIS, they also found abnormal methylation in asymptomatic healthy women who later developed breast cancer [285]. This was the first report to indicate the value of DNA methylation as a possible marker for the early detection of breast cancer and these data was further confirmed and expanded [286-288]. Our group recently was the first to perform a large-scale epigenotyping study showing *ZNF217* plasma methylation to be associated with breast cancer risk [163]. Prognostic DNA methylation markers have also been suggested by examining both serum and breast cancer

tissue. *PITX2*, *RASSF1A* and *APC* were the most frequently detected genes. Single gene loci as well as gene panels showed an association of methylation status with disease free and overall survival/mortality as well as early distant recurrence and lymph node metastasis. Finally, only a limited number of DNA methylation markers for breast cancer predicting and monitoring adjuvant treatment have been identified [116].

Regardless of the several studies no single identified marker has made the transition to the clinic. In order to improve early diagnostic and risk prediction strategies better models are needed to investigate early stage disease. In addition, more specific and sensitive markers need to be identified with studies stratifying their analysis based on the different types of breast cancer, e. g. based on the hormone sensitivity of the tumour as it has been done with genetic analysis [51]. Finally, the issue of intra-tumour heterogeneity when tumour tissues are analysed needs to be addressed.

Table 2-3: Diagnostic, prognostic and risk prediction DNA methylation biomarkers for breast cancer in different types of tissue.

Genes Identified	Tissue Analyzed	No of samples	Significant Findings	Ref.
ZNF217	Peripheral blood cell DNA	1083	Association with breast cancer risk	163
TMS1, BRCA1, ERα,PRB	Tumour tissue, normal tissue, serum	50	Potential diagnostic markers	289
$p16^{INK4A}$, $p14^{ARF}$, Cyclin D2, Slit2	Serum	36	Potential diagnostic markers	290
RASSF1A, APC, DAPK	Tumour tissue and paired preoperative serum DNA	34	Potential diagnostic markers being associated with the disease	291
GSTP1, RARß2, $p16^{INk4a}$; $p14^{ARF}$; RASSF1A, DAPK	Tumour tissue, normal breast tissue, nipple aspirate fluid	22	Potential diagnostic markers being associated with the disease	288
CCND2, RASSF1A, APC, HIN1	Needle aspirate washings	training set:109 test set:78 validation set: 45	Potential diagnostic markers being associated with the disease	292
APC,RASSF1A	Serum	122	Association with disease-free and overall survival	293
ESR1, APC, HSD17B4, HIC1, RASSF1A	Serum	training set: 24 test set: 62	Association with overall survival ion women with no adjuvant systemic therapy	293
Histone modifications (including methylation changes): H3K4me2, H4K20me3, H4R3me2 (lysine methylation)	Tumour tissue	880	Association with overall survival and tumour phenotypes	294
Kallikrein 10 (*KLK 10*), Cystatin M (CST6)	Tumour tissue	test set: 35 validation set: 93	Association with disease-free interval and overall survival	295, 296

Genes Identified	Tissue Analyzed	No of samples	Significant Findings	Ref.
RASSF1A	Tumour tissue	test set:35 validation set: 93	Association with disease-free interval	297
PITX2	Tumour tissue	241	Association with distant recurrence, disease free survival and overall survival	298
BRCA1, p16	Serum	122	Association with overall survival	299
SFRP5	Tumour tissue	133	Association with overall survival	300
PITX2	Tumour tissue	412	Association with early distant metastasis and poor overall survival	301
PITX2	Hormone receptor-positive tumour tissue	test set: 109 validation set: 236	Association with distant recurrence/metastasis	302
ID4	Tumour tissue	170	Association with recurrence free survival and lymph node metastasis	303
RASSF1A	Serum	148	Marker for monitoring of efficacy of adjuvant tamoxifen treatment	304
NEUROD1	Tumour tissue and serum	74 (tumour tissue) 44 (pre-treatment core biopsies), 107 (serum)	Marker for monitoring adjuvant treatment; Association with RFS and overall survival	305
PSAT1	Tumour tissue (steroid hormone receptor–positive)	200	Association with tamoxifen therapy response and progression free survival	306
ESR1, CYP1B1	Tumour tissue	148	Association with tamoxifen therapy response and disease free survival	307

Adopted from Jones et al, 2010 [113]

3 IDENTIFICATION OF BREAST CANCER CASES - CANCER REGISTRY VS SELF-REPORTING

In this thesis breast cancer cases were identified from among 189,046 women from England and Wales participating in a national screening study (UKCTOCS). Data on breast cancer diagnosis was available for two data sources – cancer registry follow-up and from self-reporting on the UKCTOCS follow-up questionnaire.

National cancer registries are found in many countries and collect comprehensive cancer information for the whole population which enables documentation of historical trends in cancer incidence / survival over long periods of time. The information is used for research, education and for planning national strategies to deliver the best cancer care to the whole population. The cancer registries work on a country-specific policy and therefore the availability of cancer data differs between countries. In the UK, the registries are divided between England and Wales, Scotland and Northern Ireland. In England and Wales, the NHS Information Centre for Health and Social Care (formerly the Office of National Statistics, ONS) provides data on cancer registrations through the NHS Cancer Registry (NHSCR) and data on death and cause of death through the Death Certification process (Medical Certificate of Cause of Death). In Scotland, this is through the Scottish Cancer Registry and the General Registry Office for death certificates, while in Northern Ireland Cancer Registry and the Central Services Agency (CSA) provide data on cancers and deaths, respectively.

All cancer registries collect information on every new diagnosis of cancer occurring in their populations. The information is acquired from a variety of sources including hospitals, cancer and treatment centres, hospices and private hospitals, cancer screening programmes, other cancer registers and death certificates, general practices and nursing homes. Processing of data involves checking the validity and completeness of the data and a complex process of clinical data linkage and consolidation. Overall, the data on cancer registrations has been shown for the most part to be reliable [308]. Major errors in International Classification of Diseases (ICD) coding are few [309] with data regarding cancer stage, grade and date of treatment being less consistent and delays occurring in the recording of the data. This suggests that even though the quality of the data may be good, improvements are needed in standardising the recording of information by clinicians [310].

Cancer registries are also often used for tracking participants in research studies where cancer diagnosis and mortality are key outcome measures. In such circumstances, completeness of information and timely notification is crucial. To compensate for possible delays in recording cancer data by cancer registries, researchers often use additional sources such as self-reporting through follow-up questionnaires or medical notes. Follow-up questionnaires are regarded as the most cost-effective way in obtaining these data [311]. However, the validity of this form of reporting is dependent on the site of cancer, with self-reported breast cancer being most accurately identified in comparison to other type of cancers such as endometrial, cervical [312] and ovarian cancer [313]. Reported sensitivity for breast cancer classification ranged between 79-98% in comparison to colon cancer, ranging from 58-89% [312, 314-320]. In addition, *in situ* cancers have been

shown to have much higher rates of misclassification by individuals than invasive cancers regardless of the site [312]. Sensitivity of self-reporting is dependent on a variety of factors such as age at diagnosis, education, previous family history and race [311, 316]. Abstraction of clinical information from medical reports obtained directly from the clinicians treating the patient are considered to be the most accurate means of collecting cancer data [321]. However this can be extremely time-consuming and expensive especially when different centers are involved [322].

The accuracy of the cancer data has major implications for research studies, especially those that include cancer risk prediction [7] and screening. Most previous studies reporting on accuracy of cancer data have used two of the three possible information sources (self-reported data on questionnaires, cancer registry records or medical notes). Only two have looked at all three sources of cancer data but analysis was limited to small subgroups within the study populations [311, 316].

The initial goal of the work in this thesis therefore was the identification of women with breast cancer via the two data sources available in the trial and further investigation through contact with the treating clinician to confirm breast cancer diagnosis and collection of histopathological information. This also provided an opportunity for breast cancer diagnosis to: **1)** explore the apparent sensitivity and positive predictive value (PPV) of the data sources (self-reported cancer data and cancer registry records versus confirmation from the treating clinician) **2)** elucidate causes of errors and discrepancies **3)** investigate the effect of time on cancer registration delays and **4)** examine the association between self-reporting and age, education and family breast cancer history.

3.1 Materials and Methods

3.1.1 Ethical Approval

The thesis protocol was developed and submitted for ethical approval. During the process, the Joint UCL/UCLH Committees on the Ethics of Human Research meeting was attended and all questions that were raised were answered. No major amendments were required. The study was approved on 22nd February 2007 (06/Q0505/102).

3.1.2 Subjects in UKCTOCS

The subjects were participants in UKCTOCS; the largest multi-centre randomised controlled trial for ovarian cancer that involves a cohort of 202,638 postmenopausal women from the general population recruited from 2001-2005. Details of the study design and screening interventions are available from Menon *et al* [78] and the trial website (www.ukctocs.org.uk) (screening continues until the end of 2011 and the primary endpoint of mortality reduction through screening will then be documented until 2014). Briefly, the trial was set up at 13 NHS trusts in England, Wales and Northern Ireland and is co-ordinated by the Gynaecological Cancer Research Centre at UCL. Women aged 50-74 were randomly invited from age/sex registers of the 27 participating Primary Care Trusts. Women who accepted the invitation were provided with written and verbal information about the trial. In addition, they viewed an information video at the recruitment interview. Written consent was obtained which included access to their medical records and use of their data/samples in future studies. Each woman filled in a baseline questionnaire regarding medical and family history (Appendix I). This included questions on previous history of any cancer (ovarian, breast, bowel, and lung), HRT use, and data on parity, hysterectomy, sterilisation operation, treatment for

infertility, contraceptive pill use. All the data was entered onto a sophisticated custom-built Trial Management System which confirmed their eligibility to participate in the trial.

3.1.3 Identification of breast cancer cases in UKCTOCS

The subjects for the purposes of this study were women residing in England and Wales identified by the cancer registries or self-reporting (through UKCTOCS FUQ) to have developed breast cancer by 2nd of February 2009 following randomisation to UKCTOCS. For these women who were initially identified through cancer registry and self-reported information was not available the UKCTOCS FUQ was sent to obtain information on cancer reporting. Women with benign conditions or *in situ* carcinomas of the breast were not included in the study subjects. Women recruited from Northern Ireland were excluded as data from the Northern Ireland cancer registry became available only in 2008 after the project had already started.

Cancer Registry (CR)

All women participating in the trial are ''flagged'' using the NHS number for cancers and deaths through the NHSCR. As a result the computerised entry of each subject at the registry is ''flagged'' so that the UKCTOCS coordinating centre at UCL can be notified of any deaths or new diagnosis/recurrence of cancer. The information is sent using the ICD and Health Related Problems Codes, 9th and 10th revision (ICD-9 and -10 Codes – two different editions of the cancer registry coding) and includes cancer site, morphology and date of diagnosis. For the purposes of this project the CR data was examined to identify breast cancer using ICD codes as listed in Table 3-1. Regular downloads from

the relevant cancer registries are received every 6 months in the trial centre. For women with *in situ* carcinoma of breast the following ICD codes are used: DO5* (ICD-10) or 233* (ICD-9). The codes include: LCIS and DCIS.

Table 3-1: **International classification of breast cancer, ICD -9, -10 codes for breast cancer (invasive malignant neoplasm of breast).**

Invasive malignant neoplasm of breast C50 (ICD-10) and C174 (ICD-9)		
Includes: connective tissue of breast ***Excludes***: skin of breast (C43.5 , C44.5)		
C50.0 and C174.0	-	Nipple and areola
C50.1 and C174.1	-	Central portion of breast
C50.2 and C174.2	-	Upper-inner quadrant of breast
C50.3 and C174.3	-	Lower-inner quadrant of breast
C50.4 and C174.4	-	Upper-outer quadrant of breast
C50.5 and C174.5	-	Lower-outer quadrant of breast
C50.6 and C174.6	-	Axillary tail of breast
C50.8 and C174.8	-	Overlapping lesion of breast
C50.9 and C174.9	-	Breast, unspecified

ICD=International classification of diseases

UKCTOCS Follow-Up Questionnaire (FUQ)

The UKCTOCS protocol included a follow-up questionnaire 3.5 years after randomisation (Appendix II). The 11-item FUQ included items on cancer diagnosis since randomisation and a specific question related to breast cancer, education, alcohol consumption, smoking status, skirt size, HRT use. Women who reported breast cancer were asked to provide the name of the treating physician (consultant), the hospital where they were treated and the year when surgery/biopsy was undertaken.

3.1.4 Confirmation of breast cancer diagnosis through the collection of clinicopathological data in a form of questionnaire

For all women who were identified to have developed breast cancer after randomisation, the diagnosis was confirmed by sending a Breast Cancer Questionnaire (BCQ) (Appendix III) specifically designed for the purposes of the study to the consultants treating the women. The 15-item BCQ included questions on site of tumour, grade, stage, histology, receptor status (ER, PR and HER2/neu), diagnosis date and treatment. The consultants had the option to provide a histopathology report if they were unwilling to complete the questionnaire. Some of the consultants provided both the questionnaire and the histopathology report. Individualised letters were sent to consultants where there was missing data on the returned BCQ (A-BCQ). Those who did not return the BCQ within four months were sent a second questionnaire (R-BCQ). A copy of the consent form was not routinely sent but was provided to consultants on request. In some cases, it was not possible to post a BCQ to the treating physician as the contact details were missing or incomplete or the only source of information was a death certificate with a breast cancer diagnosis. In such cases, the UKCTOCS research nurse at the regional centre where the woman was registered was asked to search the medical notes for a histopathology report.

3.1.5 Socio-demographic characteristics of study subjects

From the baseline recruitment questionnaire and FUQ socio-demographic characteristics were collected as mentioned above. The following factors were analysed to investigate whether self-reporting is dependent on them: (1) race (*white*/*non white*) (2) breast cancer family history (*no* and *yes*, including first and second degree relatives such as mother, sister, grandmother, granddaughter and

aunt) as recorded by the women in the baseline recruitment questionnaire (3) Education (*high*: university/university college, *low*: college: A – and O- level, qualifications such as clerical and commercial e.g. hairdressing, and *none*: either not reporting anything or reporting that they did not have any of the above education, as recorded in the FUQ.

3.1.6 Database development and data storage

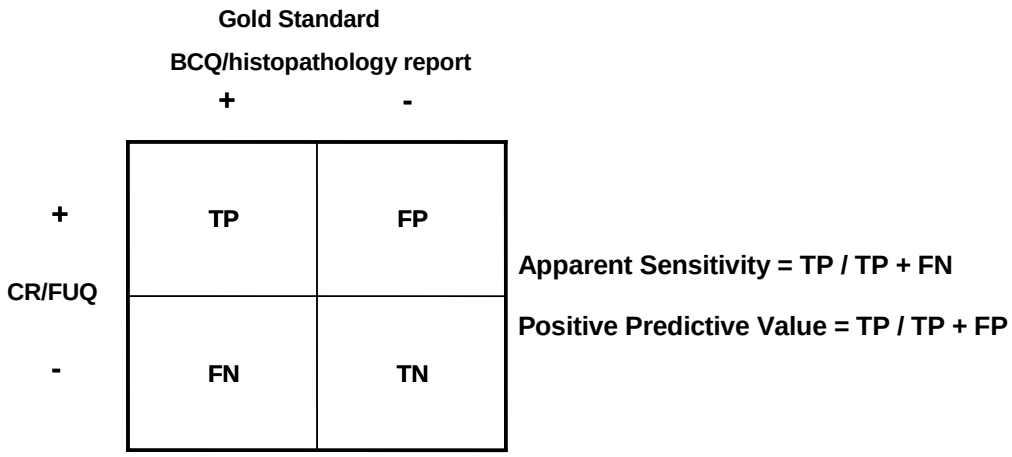
As part of the work undertaken in the course of this thesis, an ACCESS database was built to enter the study data. It had two main tables - one holding general information on women identified to have breast cancer (Breast Cancer Table) and the second with all the data collected from the BCQs/histopathology reports (Clinical Data Table).

3.1.7 Data analysis

When data collection was complete, a flow diagram along the lines of CONSORT flow chart was prepared which included the outcome in each of the women initially identified to have breast cancer. All women for whom it was possible to obtain data from the three sources (CR, FUQ and BCQ/histopathology report) were included in the final study subjects in order to investigate the sensitivity and PPV of CR and self-reporting and to identify the eligible cases for the study described in chapter 4. Baseline characteristics and histopathological information were calculated using descriptive statistics. If both breast cancer and *in situ* carcinoma of breast was reported in the same woman, the breast cancer diagnosis was used for comparisons.

For the secondary objectives, analysis was undertaken comparing CR and FUQ with the gold standard - BCQ/histopathology report. Misclassifications were identified for CR and FUQ individually. The true positive (TP), false positive (FP), false negative (FN) and true negative (TN) were assessed and the apparent sensitivity and positive predictive value (PPV) of each data source was calculated as shown in Figure 3-1. Apparent sensitivity was used as it was not possible in this study plan to identify women with breast cancer who did not self-report breast cancer or had cancer registration (true negatives TN) since their physicians were not asked to provide a BCQ/histopathology report.

Figure 3-1: Description of performance characteristics used for data analysis. Calculation of apparent sensitivity and positive predictive value (PPV) for cancer registry (CR) and follow-up questionnaire (FUQ) data using the breast cancer questionnaire (BCQ)/ histopathology report as the gold standard.



BCQ=breast cancer questionnaire; CR=cancer registry; FN=false negative; FP=false positive; FUQ=follow-up questionnaire; PPV=positive predictive value; TN=true negative; TP=true positive

In most situations where large numbers of women with breast cancer need to be identified, it is not possible to obtain confirmation through the physician. To address this issue, combining CR and FUQ data was explored using the following rules - a. breast cancer case is correctly reported if both sources concurred for breast cancer diagnosis and b. breast cancer case is correctly reported if either source (CR or FUQ) reported breast cancer diagnosis. Fisher's test was used to compare sensitivities and PPVs of CR and FUQ.

The effect of time on cancer registration delays was assessed by looking at the completeness of relevant cancer registrations according to year of diagnosis and time from diagnosis to CR notification. The effect on apparent sensitivity and PPV of age at FUQ, race, education and family history was investigated. Apparent sensitivity was modelled using logistic regression with the above characteristics as the independent variables, and using only those cases where the BCQ/histopathology report confirmed breast cancer. All four variables were suitably categorised before modelling and from the regression, the respective odds ratio and significance levels were estimated, given the other variables' presence in the equation. PPV was also modelled in exactly the same way, using only those cases where women self-reported positively. Analysis was carried out using a computer assisted program-SPSS version 12.0.1, Chicago, IL.

3.2 Results

3.2.1 Identification of breast cancer cases in UKCTOCS

Of 189,046 women recruited into the trial from England and Wales between 2001 and 2005, 2629 women were identified as having breast cancer post randomisation by 2nd February 2009 either by cancer registry or self-reporting. It is to be noted that this was heavily skewed towards initial identification through CR (total number of women identified through CR were 2475) as UKCTOCS FUQ had not been sent to most women when this study commenced (for 460 women who reported breast cancer and there was also a cancer registration – these number of women only were used for the purposes of the analysis since our aim was to obtain information from all three sources for as many women within the UKCTOCS cohort). In addition to the three sources, in 10 women breast cancer was identified as a result of ovarian cancer screening in UKCTOCS which resulted in raised serum CA125 levels.

Table 3-2: Primary source of breast cancer notification in UKCTOCS

Primary source of notification for breast cancer	
CR	
ONS	2015
Death certificates	154
CR and Self-Reporting	
FUQ	460*
*10 women identified during screening	
Total No of women	**2629**

CR=Cancer registry; FUQ=follow-up questionnaire; ONS=Office of national statistics

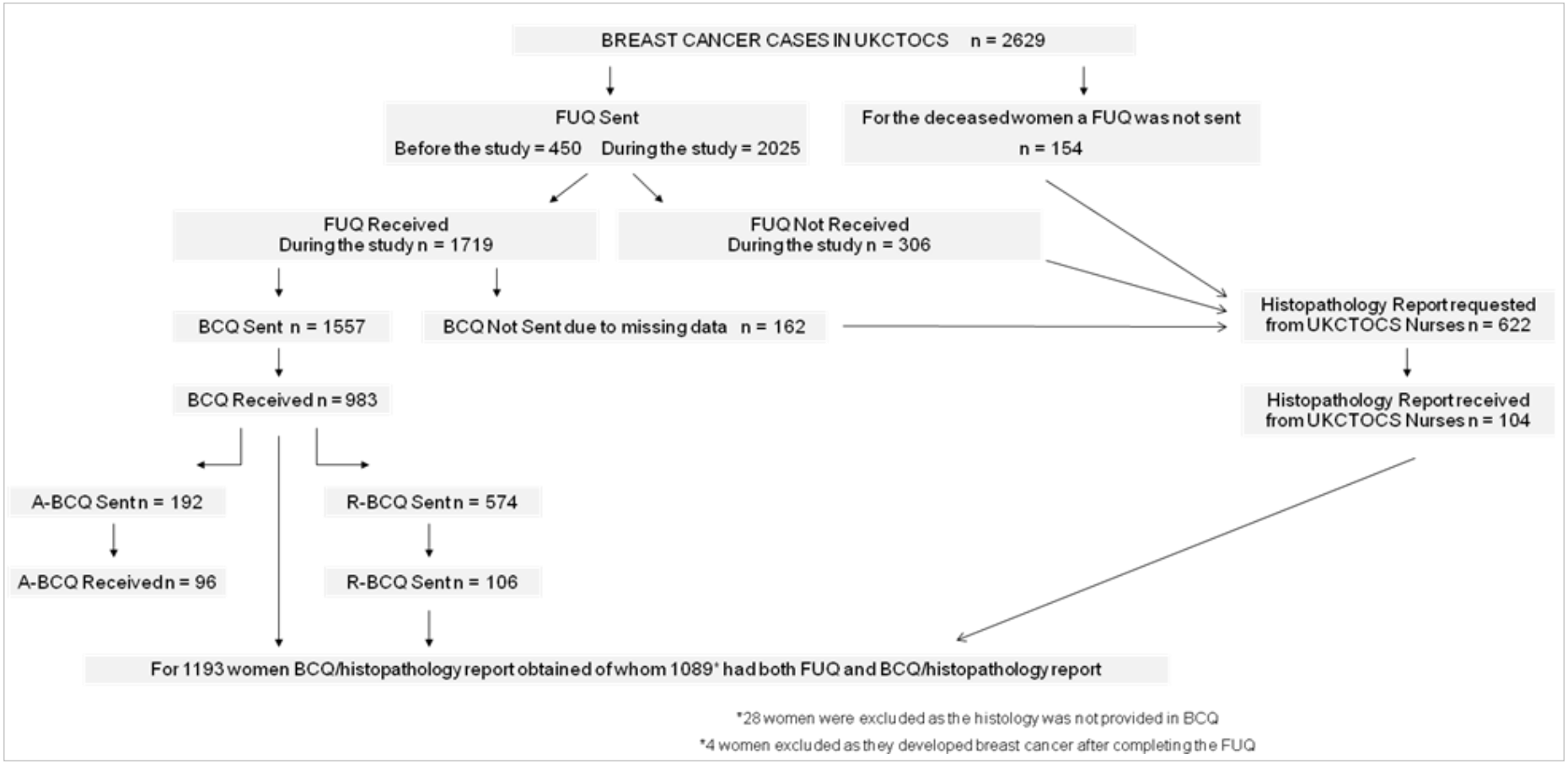
The 2015 women identified through CR and the 10 cases identified during screening, were sent a UKCTOCS FUQ as shown in Figure 3-2. It was not possible to send FUQ to the 154 women for whom breast cancer was first identified through the death certificate. The response rate for FUQs sent during

this study was 84.9% (1719 of 2025). The overall FUQ response rate for UKCTOCS at the time of this study was 74.6% (115396/154590) for England and Wales. BCQ was sent to 1557 women (Figure 3-2) with a response rate at 63.1% and R-BCQs were sent to 574 with a response rate at 18.5%. The overall response rate was therefore 70% (1089 of 1557). To obtain missing data the A-BCQs were sent to 192 consultants, with a response rate of 50%. BCQs could not be sent to the physicians for 622 women who did not provide contact details of their consultant. The UKCTOCS trial nurses were contacted in order to obtain the histopathology reports for these women. 104 reports were obtained giving a response rate at 16.7%.

For 1089 women data from all three sources was obtained. 32 women had to be excluded from the analysis as complete histological information was missing in 28 and 4 women were diagnosed after completion of the FUQ. Therefore, the final number of eligible women in this study was 1057. In 23 women where CR gave notification of both breast cancer and *in situ* carcinoma of breast, the breast cancer registration was used in the analysis. In an additional 95 women CR reported a cancer other than breast which was not taken into consideration for the purposes of this analysis.

Figure 3-2: Diagram showing how the study subjects were identified.

(Collection of three sources; cancer registry, self-report through UKCTOCS follow-up questionnaire and breast cancer questionnaire).



A-BCQ= additional breast cancer questionnaire; BCQ= breast cancer questionnaire; CR=cancer registry; FUQ= follow-up questionnaire; R-BCQ= reminder breast cancer questionnaire

3.2.2 Distribution and frequency of socio-demographic characteristics of the study subjects

Women were reported with breast cancer between 2001 and 2008 with nearly 30% being identified in 2005. The median age of the women at breast cancer diagnosis was 62 years (range 50-78 years). The median age of the women at self-reporting (FUQ) was 64 years (range 52-80 years). 97.9% of the women were white, 33.1% were university graduates, and 27.6% had at least one 1st / 2nd degree relative with breast cancer history (Table 3-3).

Table 3-3: Distribution and frequency of sociodemographic characteristics of the study subjects. (N=1057)

Characteristics	No of Women	%
Age at diagnosis		
50-64	665	62.9
65-80	392	36.9
Age at FUQ		
50-64	557	52.5
65-80	500	47.1
Race		
White	1039	97.9
Non-White	14	1.3
Unknown	4	0.4
Education		
None	331	31.2
Low	372	35.1
High	354	33.1
Year of breast cancer diagnosis		
2001	3	0.3
2002	37	3.5
2003	117	11
2004	250	23.6
2005	313	29.5
2006	257	24.2
2007	79	7.4
2008	1	0.1
Breast cancer family history		
Yes	293	27.6
No	764	72

3.2.3 Performance characteristics for CR and self-reporting through the FUQ

On comparing CR with the BCQ/histopathology report (gold standard), 30 (3.2%) FP cases were identified having a breast cancer registration code despite not having breast cancer according to their physician (BCQ/histopathology report). 29 had DCIS on BCQ/histopathology report and 6 of these women had a breast registration code of carcinoma *in situ* in addition to their breast cancer registration code. The remaining one FP had atypical ductal hyperplasia. There were 47 FN cases; this included one woman who had a neck cancer registration 2 years and 3 months prior to breast cancer diagnosis. Seven (0.7%) of the 47 FN had an *in situ* carcinoma of breast registration code instead of a breast cancer registration code and 2 of these had DCIS as well as breast cancer on BCQ/histopathology report. Forty (4.3%) of the 47 FN cases were not registered since the last CR follow-up (2nd February 2009) (Figure 3-3 A and Table 3-4) and all of them were diagnosed with an invasive breast cancer according to their physician (BCQ/histopathology record). Overall, on BCQ/histopathology report, 112 women had DCIS and 1 LCIS and 3 had benign conditions of which 74 of the DCIS and 2 of the benign breast conditions had a corresponding *in situ* carcinoma registration code. The apparent sensitivity of CR was 95.0% (93.4 to 96.2) and PPV was 96.8% (95.3 to 97.8) (Figure 3-3 C).

On comparing the FUQ with BCQ/histopathology report (gold standard), there were 116 (12.3%) women (FP) who self-reported breast cancer despite having an *in situ carcinoma* or benign conditions (112 women had a diagnosis of DCIS, 1 LCIS, and 3 benign breast conditions; atypical ductal hyperplasia, fibrocystic changes and non invasive papillary lesion) as confirmed by their physicians on

the BCQ/histopathology report. Moreover, 35 (3.3%) women (FN) did not self-report breast cancer on the FUQ which was completed a median 4 months (range 53 days up to 4.7 years, Interquartile range: 1.4 years) after breast cancer diagnosis (Figure 3-3 B and Table 3-4). The apparent sensitivity of self-reporting on the FUQ was 96.3% (94.9 to 97.3) and PPV was 88.7% (86.5 to 90.5) (Figure 3-3 C).

Table 3-4 summarises the discrepancies/errors identified by comparing the two sources (CR and self-reporting through the FUQ) to BCQ/histopathology report (gold standard). Out of 941 (89.0%) women with confirmed breast cancer diagnosis on histopathology, both CR and self-reporting concurred in 859 (69.5%) women diagnosed with breast cancer. 77 (8.2%) women would have been missed if CR alone was used and 151 (16.0%) women would have been missed if FUQ alone was used. When the rule that both sources (CR and FUQ) need to concur for breast cancer diagnosis was applied, there were 30 (3.2%) women who would have been falsely identified as breast cancer cases. When the rule that breast cancer case is correctly reported if either source (CR or FUQ) reported breast cancer diagnosis was applied, there were 168 women (17.9%) who would have been misclassified or not reported/registered. The lowest rate of misclassifications (3.2%) was observed when breast cancer diagnosis was confirmed by both sources; CR and self-reporting through FUQ.

Figure 3-3: **Performance characteristics for cancer registry and UKCTOCS follow-up questionnaire.**

Numbers of true positives (TP), false positives (FP), false negatives (FN) and true positives (TP) for breast cancer cases identified within UKCTOCS. Comparison with gold standard (BQC/histopathology) of (A) CR and (B) self-reporting through FUQ. C) Calculation of sensitivity and positive predictive value (PPV).

(A)

CR \ BCQ (Gold Standard)	+	-
+	894	30
-	47	86

(B)

FUQ \ BCQ (Gold Standard)	+	-
+	906	116
-	35	0

(C)

Performance Characteristics	CR	FUQ	p-value
	BCQ (Gold Standard)		
% Apparent Sensitivity (95% CI)	95.0 (93.4 to 96.2)	96.3 (94.9 to 97.3)	**0.2140**
% PPV (95%CI)	96.8 (95.3 to 97.8)	88.7 (86.5 to 90.5)	**<0.0001**

BCQ=breast cancer questionnaire; CR=cancer registry; FUQ=follow-up questionnaire; PPV=positive predictive value

Table 3-4: Identified misclassifications/errors by comparing all three different sources and their causes. The error is dependent on the data source and how it is interpreted. % of misclassifications were calculated based on the total number of women confirmed with breast cancer diagnosis (N=941).

Cause of misclassification	Data source and interpretation			
	FUQ	**CR**	**CR and FUQ (both need to concur for BC diagnosis)**	**CR and FUQ (BC diagnosis if either report BC diagnosis)**
DCIS or benign condition misclassified as BC	116 (12.3%)	30 (3.2%)	30 (3.2%)	86 (9.1%)
BC misclassified as DCIS or benign condition	0 (0%)	7 (0.7%)	0 (0%)	7 (0.7%)
BC not reported/registered	35 (3.7%)	40 (4.3%)	0 (0%)	75 (8.0%)
Total No (%) of missed study subjects (either misclassified or not reported/registered)	151 (16.0%)	77 (8.2%)	30 (3.2%)	168 (17.9%)

BC=breast cancer; CR=cancer registry; DCIS=ductal carcinoma *in situ*; FUQ=follow-up questionnaire

3.2.4 Distribution and frequency of the clinicopathological characteristics of the confirmed breast cancer cases within UKCTOCS

Histopathological data and treatment information of the confirmed breast cancer cases (N=941) is provided in Table 3-5. Majority of the women had breast cancer on their left breast. 44.7% of the women were diagnosed with Grade II and 41.2% had a Stage 1 tumour. Most of the women (74%) did not have a lymph node metastasis and 1.8% of women had a metastasis in a distant organ. The highest percentage of women was diagnosed with IDC (70.1%). 73% of the women were diagnosed with ER positive breast cancer, 38% with PR positive and 11% with HER2 positive. Regarding treatment, 61.1% of the women had radiotherapy, 50.4% of the women had WLE, and 48.9% had ANC. The most common systemic therapy used was tamoxifen.

Table 3-5: Clinicopathological characteristics of the confirmed breast cancer cases within the UKCTOCS cohort. (N=941)

Clinicopathological Data		No of women	%
Primary tumour site	Left	473	50.3
	Right	413	43.9
	Bilateral	31	3.3
	Missing	24	2.6
Grade	I	210	22.3
	II	421	44.7
	III	234	24.9
	Other	32	3.4
	Missing	44	4.7
Stage	1	388	41.2
	2	207	22.0
	3	30	3.2
	4 or 5	3	0.3
	Combination of 1/2/3	36	3.8
	Missing	279	29.6
Lymph node involvement	Yes	245	26.0
	No	696	74.0
Metastasis to distant organ	Yes	17	1.8
	No	924	98.2

Clinicopathological Data		No of women	%
Histology	IDC	602	64.0
	IDC&DCIS	57	6.1
	ILC	110	11.7
	ITC	3	0.3
	IDC&ILC (Mixed)	32	3.4
	Other	137	14.6
ER	Positive	687	73.0
	Negative	143	15.2
	Not done	5	0.5
	Borderline	2	0.2
	Missing	106	11.3
PR	Positive	360	38.3
	Negative	196	20.8
	Not done	11	1.2
	Borderline	8	0.9
	Missing	366	38.9
HER2	Positive	104	11.1
	Negative	296	31.5
	Not done	47	5
	Missing	498	52.9
Radiotherapy	Yes	575	61.1
	No	366	38.9
Surgery-Breast	WLE	474	50.4
	Simple Mastectomy	297	31.6
	Radical Mastectomy	29	3.1
	Lumpectomy	41	4.4
	None	4	0.4
	Combination of the above	49	5.2
	Missing	47	5.0
Surgery-Nodes	SLN	185	19.7
	ANC	460	48.9
	ANS	145	15.4
	No dissection	16	1.7
	SLN and ANC	24	2.6
	SLN and ANS	35	3.7
	Missing	76	8.1
Systemic Therapy	Tamoxifen	318	33.8
	Aromatase Inhibitors	130	13.8
	Anthracyclines	37	3.9
	Non Anthracyclines	6	0.6
	Herceptin	6	0.6
	Combination of the above	231	30
	Missing	163	17.3

ANC=axillary node clearance; DCIS=ductal carcinoma *in situ*; ER=oestrogen receptor; IDC=invasive ductal carcinoma; ILC=invasive lobular carcinoma; ITC=invasive tubular carcinoma; PR=progesterone receptor; WLE=wide local excision

3.2.5 Cancer registration delays

As in some instances, there are delays in CRs, it was important to investigate whether this might account for the lack of cancer registration in the 47 women who were confirmed on BCQ to have breast cancer but did not have a breast cancer registration on 2nd February 2009. The year of diagnosis of breast cancer in these 47 women were 2003 in 3 (2.6%), 2004 in 6 (2.4%), 2005 in 7 (2.2%), 2006 in 8 (3.1%) and 2007 in 23 (29.1%) (Figure 3-4). Between 2001-2002, there were no women diagnosed with breast cancer not having breast cancer registration code. The highest percentage of women (29.1%) being missed on CR were those diagnosed within the year 2007.

Figure 3-5 which examines time for diagnosis shows that the majority of women without a cancer registration were those diagnosed 1 to 2 years prior to the date of last CR follow-up (2nd February 2009). For all women with breast cancer in the study subjects who were identified from the two data sources (CR and FUQ) between 2001-2008 and compared to BCQ/histopathology report, 7, 18, 8, 6, 7 and 1 were not registered after 1 up to 6 years respectively.

Figure 3-4: % of women without a breast cancer registration code based on the last cancer registry follow-up (2nd February 2009) in relation to the number of years prior to diagnosis.

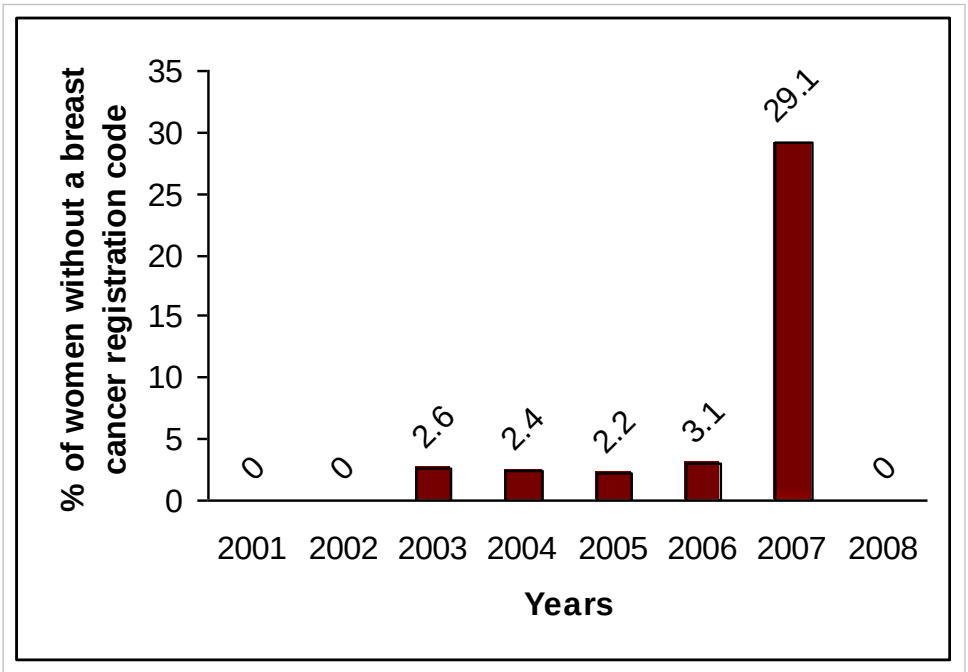
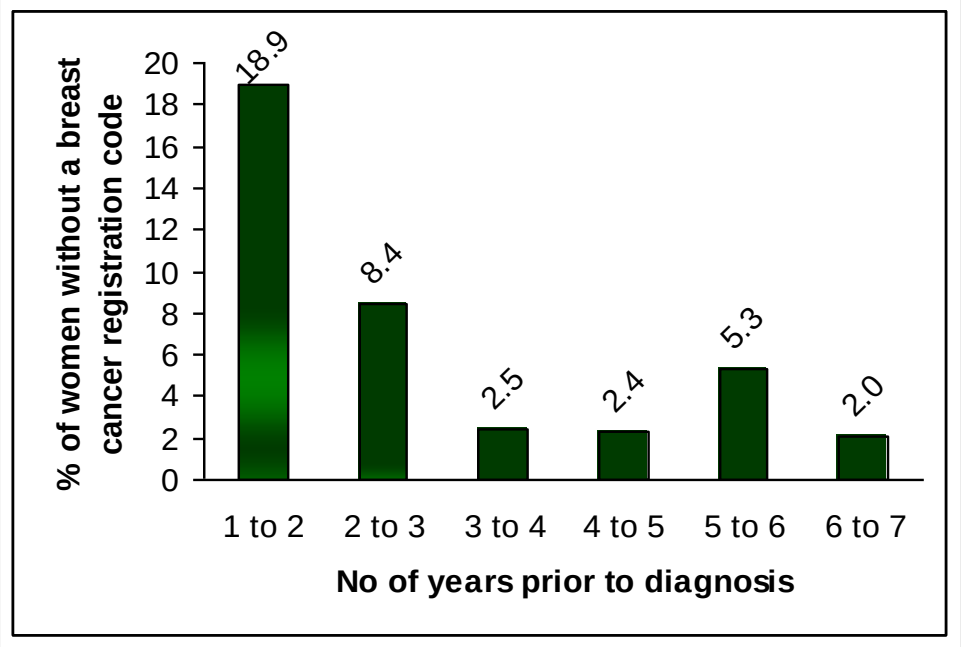


Figure 3-5: % of women without a breast cancer registration code per year based on the last cancer registry follow-up (2nd February 2009).



3.2.6 Apparent sensitivity and PPV of self-reporting based on characteristics of the study subjects

In order to investigate whether self-reporting is affected by factors such as age, race, education and breast cancer family history, we calculated the apparent sensitivity and PPV in relation to the above mentioned study characteristics (Table 3-6). Education was the most significant determinant of apparent sensitivity and borderline significant for PPV, with more educated women correctly reporting their breast cancer diagnosis in comparison to women with no education. Breast cancer family history was a significant determinant of apparent sensitivity but not for PPV, with women having a relative with breast cancer compared to respondents who did not have any relatives with breast cancer under-reporting their breast cancer diagnosis. Age was a significant determinant for PPV but not for apparent sensitivity, with women <65 in comparison to women >65 over-reporting their breast cancer diagnosis. Both apparent sensitivity and PPV did not differ by race.

Table 3-6: Characteristics of the study women as determinants of apparent sensitivity and positive predictive value. The respective odds ratio and significance levels were estimated.

Variable	Apparent Sensitivity	OR	95%CI	p-value	PPV	OR	95%CI	p-value
Race								
White	96.38	1.00			87.55	1.00		
Non-White	100.00	(not measurable)	0.00	-	92.85	1.64	0.21 to 12.82	0.637
Unknown								
Age at FUQ								
50-64	96.65	1.00			85.55	1.00		
65-80	96.22	0.72	0.33 to 1.53	0.394	90.02	1.57	1.10 to 2.32	**0.022**
Family Breast Cancer History								
No	97.01	1.00			87.5	1.00		
Yes	94.98	0.35	0.16 to 0.74	**0.006**	88.1	0.87	0.56 to 1.35	0.534
Education								
None	90.17	1.00			84.81	1.00		
Low	99.10	10.85	3.72 to 31.63	**0.000**	90.24	1.64	0.99 to 2.72	0.057
High	99.34	24.42	3.25 to 183.20	**0.002**	87.42	1.53	0.99 to 2.34	0.055
Total	**96.3**				**88.7**			

CI=confidence interval; FUQ=follow-up questionnaire, OR=odds ratio; PPV=positive predictive value

3.3 Discussion

As a result of this study, 2629 women were identified on either CR or FUQ update on 2nd February 2009 to have breast cancer following recruitment to UKCTOCS. Using the BCQ, invasive breast cancer diagnosis was confirmed in 941 women. These women formed the cohort used for the identification of eligible cases in Chapter 4.

The distribution of Stage (TNM), histology and treatment was similar to that reported for women diagnosed with breast cancer in cohorts described in England, in 2007 and 2009 [323, 324] and to be representative of any breast cancer cohort in Europe compared to a recent publication by ONCOPOOL – a European database that includes 16,944 breast cancer cases [325]. In general, it has been shown that on average 80% of breast cancers are IDC and 5-15% are ILC, 70% are ER positive, 25% are HER2 positive and 5% of the breast cancers will metastasize in a distant organ and almost 30% will have a nodal status positive [11, 14]. Regarding treatment, majority of the study women had radiotherapy, WLE and hormonal therapy. This observation comes in agreement with previous cohorts studied in England [323, 324].

Based on the data collected the apparent sensitivity and PPV of CR and FUQ were calculated making this study the first in England and Wales to examine the performance characteristics of both self-reporting and CR for breast cancer diagnosis in comparison to a report from the treating physician. A high sensitivity was observed for both FUQ and CR but PPV was significantly lower for FUQ compared to CR. For breast cancer, using national CR data for England and Wales would result in an error rate of 8.2%. However half (4.3%)

of this is related to the two year time delay in registration so that allowing for this, error rates could be reduced to 3.9%. Self-reported data is not associated with time delays but is dependent on age, family history and education. Misclassifications would be in the range of 16% (3.7% of women may not report breast cancer and 12.3% may self-report breast cancer despite having only *in situ* carcinoma or benign condition). If confirmation from the physician is not available then the most accurate source of information would involve combining CR and self-reporting data using the rule that both must concur if breast cancer diagnosis is to be confirmed. This is associated with the lowest rate (3.2%) of misclassifications.

One of the main advantages of this study is that the consultants responsible for treating the women were contacted to obtain data regarding breast cancer diagnosis which could be used as the gold standard. Physicians were contacted on multiple occasions to obtain as complete data as possible. Another is the size of the study. Except for eight women who refused consent to CR 'flagging', we were able to 'flag' all 189,038 women taking part in UKCTOCS from England and Wales. This was due to having accurate NHS numbers of all women prior to invitation to the trial as a result of electronic transfer from Primary Care age-sex registers [78]. In addition, by the time, this study was undertaken, 154,590 of the 189,038 women had been sent questionnaires to gather data on self-reporting. Moreover, as the trial invited over 1.2 million women, aged 50 to 74, randomly selected from England and Wales, it possible to extrapolate the findings to women from the general population belonging to this age group. The high response rates of FUQ (85%) and BCQ (70%) add to the strength of the study. The latter is especially notable as busy consultants who were not trial

collaborators completed the questionnaire. The response rate of second requests for information (18.5%) was low and suggests that there is little to be gained by contacting consultants who did not provide information initially. For future studies, it would be also useful coming in contact with the consultants through telephone and investigate whether such approach could improve the response rates.

One limitation of the study was that the treating physicians of every woman in the cohort were not contacted to identify women with breast cancer due to the significant resource issues related to contacting making it only possible to determine apparent sensitivity and PPV of self-reporting (FUQ) and CR. Other parameters such as specificity and negative predictive value could not be accurately estimated. An earlier study suggested that individuals usually tend to under-report rather than over-report breast cancer history [319]. In this study though, there were more women over-reporting their diagnosis (self-reporting *in situ* or benign condition as breast cancer).

For future studies, it would be possible to also come in contact with General Practitioners (GPs) in order to investigate how accurate these data could be in relation to the other sources. Previous studies have shown conflicting results about the validity of information being obtained from GPs. A study comparing the Northern and Yorkshire CR and Information Service with GP data, obtained by 5 practices, on cancer diagnosis, reported that GP responses were not able to identify the majority of patients diagnosed with a cancer. There was a poor level of completeness (29.4%) and correctness (65.6%) when compared with CR [326]. However, this is in contrast to a comparison of GP Research Database to ONS in England [327].

The national CR data failed to identify 5% (47/941) of women with a confirmed breast cancer diagnosis. 40 (4.3%) women had no cancer registration and 7 (0.7%) had *in situ* carcinoma of breast registration. The latter 7 women would have been classified as over-reporting in studies using only CR and FUQ. Confirmation of diagnosis in our study eliminated this bias. Brewster *et al* investigating the Scottish CR versus data from 5 independent clinical trial databases reported 0.3% of women being incorrectly classified as breast cancer when they had carcinoma *in situ* of the breast [320]. When delays in national cancer registration were investigated, it was seen that 18.9% (7/40) of women who did not have a cancer registration were within 1-2 years of their diagnosis. Sensitivity of CR has improved over the last decades (from 72% in 1987 to 95% in this report) and it is likely that further improvements in the recording of cancer data by the regional cancer registries will result in complete data as that seen in the Scandinavian countries [328] (Table 3-7). In order to meet the growing demand for timely and accurate data about cancer registration, it has been suggested that CRs should be provided with additional support so that there would be an enhancement in their capability to rapidly ascertain cancer cases [329].

The apparent sensitivity of 95.0% of CR in England and Wales reported in this study is comparable to the most recent report of 98.0% by Brewster *et al* investigating the Scottish CR. [320]. The rates are also similar to Gathani and his colleagues who reported a sensitivity of 96% for breast cancer diagnosis when CR and the National Health Service Breast Screening Program were compared in the largest study so far in England including more than 5,000 breast cancer cases [330]. In general though there are not many studies in England on validation of CR data and most of them include only a small number of breast cancer cases

(Table 3-7). When countries are compared on data regarding breast cancer diagnosis, Denmark has the highest rates (99%) of complete CR records [328].

In our study, the apparent sensitivity and PPV of self-reported breast cancer diagnosis through the UKCTOCS FUQ was 96.8%. This is comparable to the high sensitivity reported by Parikh-Patel *et al*, 2003 (96.0% for both invasive and *in situ*, 98.1% for invasive and 87.8% for *in situ* carcinoma) [312] and that by Abraham *et al*, 2009 who reported a sensitivity of 96.9% for breast cancer diagnosis and 90.2% for DCIS [311] (Table 3-7). It is not possible in our study to comment on DCIS as the primary aim was to identify breast cancer cases. What needs to be pointed though is that 12.3% of women over-reported their diagnosis with a PPV value for the FUQ at 88.2%.

Previously, it has been shown that a variety of factors affect self-reporting including age, sex, education level and family history of the investigated disease [311, 316]. In this study, we examined race, age at completion of FUQ, education and family history of breast cancer. Even though analysis was performed for race, real conclusions cannot be withdrawn as the majority of the volunteer women in UKCTOCS and our study participants were White.

Age at completion of FUQ did not make any difference in the apparent sensitivity of reporting but it did significantly correlate with PPV. Younger women were more likely to give a false positive history than the older respondents. Previous studies have shown age to have an effect on under-reporting but also over-reporting [312, 315, 316, 318, 319] . It is to be noted that in majority of the cases the FUQ was sent 2 years following diagnosis and the two ages (at diagnosis and at completion FUQ)

highly correlated and therefore only the latter was included in the regression analysis model. As reported by others [311, 315-317], women who are less educated have a greater possibility to falsely self-report cancer diagnosis. Our observation was similar with women who had been to college or university having less false positive when self-reporting breast cancer diagnosis in comparison to women who did not report anything on the FUQ regarding their education or reported no education. It has been suggested that women with a family history of the disease are better responders when they are asked about their breast cancer diagnosis [311]. In this study, though, the opposite observation was made as women having relatives with breast cancer history significantly under-reported their diagnosis. It is unclear as to what the explanation for this might be.

Furthermore, it is worth mentioning that in general on self-reporting, the sensitivity for breast cancer is better in comparison to other cancers with breast cancer having highest percentage agreement, followed by bowel and then lung cancer [319]. In 1993, a US study showed that best rates of confirmation were for breast, bladder, prostate and uterine cancer but that the rates decreased in the closely related sites, such as colon and rectum [314] indicating that use of self-reporting for more diagnostically complex diseases may require additional confirmation.

In conclusion, the data in this study informs researchers who plan epidemiological studies or trials to rely on CR as in general; the percentage of misclassification is low especially if time delays are taken into consideration. While self-reporting using postal questionnaires is another good source of cancer data, several factors such as education, age and family history need to be taken into account. Confirmation of the data by checking medical notes would be ideal

as misclassifications by both sources may occur. In the absence of the latter, the most accurate source of information involves combining CR and self-reporting data using the rule that both must concur if breast cancer is to be confirmed. This would result in around 3% misclassifications but this need to be balanced against the cost and time to researchers to collect data from medical notes.

Table 3-7: Summary of previous studies indicating the percentage agreement/sensitivity of different sources for breast cancer cases identification.

Summary of previous studies investigating completeness of breast cancer diagnosis						
Source of Identification	**Author**	**Year**	**Period covered**	**Country**	**No of participants**	**(%) agreement/sensitivity**
Self-Report vs CR	Dominguez [319]	2007	1980-1981	United States	2624	92.1
	Manjer [318]	2004	1991-1996	Sweden	170	97
	Parikh-Patel [312]	2003	1995-1996	United States	2596	98.1
	Desai [317]	2001	1981-1982	United States	64	79.2
	Bergmann [316]	1998	1992-1993	United States	995	91
	Schrijvers [315]	1994	1991	United States	85	84
Self-Report vs Medical Record	Paganini-Hill [314]	1993	1983	United States	271	90
			1985	United States	148	45
CR vs Medical Record	Brewster [320]	2008	1978-2000	Scotland	2621	98.2
	Stotter [331]	2000	1997	England	599	89
	Villard-Mackintosh [332]	1988	1968-1985	England	150	92
	Hunt and Coleman [333]	1987	1985	England	50	72
	Jensen [328]	2002	1983-1989	Denmark	2062	99
CR vs Breast Cancer Screening Program	Gathani [330]	2005	1996-2000	England	5684	96
Self-Report vs CR and Medical Record	Abraham [311]	2009	1996-2006	United States	24631	96.9

CR=cancer registry

4 HORMONAL EFFECT IN BREAST CANCER

4.1 Introduction

Sex steroid hormones are known to be crucially involved in breast carcinogenesis and are known to increase breast cancer risk. As discussed in the literature review it has become apparent that factors which are surrogates for long term sex steroid exposure such as reproductive factors (age at menarche, first birth, parity and menopause) and breast size are associated with breast cancer risk [334] as well as several anthropometric factors such as height, weight, weight changes, BMI, fat deposition, all of which contribute to changes in sex steroid levels [185]. This has led to the hypothesis that circulating sex steroid levels can predict breast cancer risk. A number of studies have been carried out in order to identify the association of serum sex steroid hormones with breast cancer [198-202, 204-208, 211, 227, 228]. The largest meta-analysis combining nine prospective studies demonstrated that postmenopausal women with serum sex steroid hormone levels in the highest quintiles have a two-fold increased risk of breast cancer [197]. In women who develop the disease, hormonal therapy plays an increasingly significant role in treatment. It is therefore imperative that we increase our understanding of how hormones interact to increase a woman's breast cancer risk.

Sex steroids exert their effects through binding to sex steroid hormone receptors. Upon binding, the receptor travels from the cytoplasm where is located in its inactive form to the nucleus where it gets dimerized and binds to HRE. This leads to activation of transcription processes and synthesis of specific messenger RNA and protein production. All of the published studies on associations of sex steroid

hormones with breast cancer risk have used conventional immunoassays to measure hormonal levels. However, in the past few years, bioactivity assays for steroid hormones have been described, enabling quantification of total sex steroid hormonal action. As a result, our group was able to provide the first evidence that ER-α and ER-β serum bioactivity (SB) are independently associated with breast cancer using samples collected at diagnosis. Women with the highest quintile of ER-α had a 2.70 fold increase in oestrogen receptor positive breast cancer risk and women with ER-β SB had a 2.31 fold increase in oestrogen receptor positive breast cancer risk [162]. Therefore, we have hypothesised that sex steroid receptor SB could predict breast cancer risk before diagnosis and provide further information on their effect in breast carcinogenesis.

To better understand the long term effect of sex steroids and bioactivity of their receptors on breast cancer risk, it is crucial to examine levels many years prior to diagnosis. In the meta-analysis high oestrogen and androgen levels more than two years before breast cancer diagnosis were found to be associated with higher breast cancer risk in comparison to levels within two years of diagnosis. This suggests that the positive associations between sex steroid hormone levels and breast cancer prior to diagnosis are more likely to be due to the effect of hormones on the development of breast cancer rather than an effect of the pre-clinical tumour on hormone metabolism [197]. Additionally, even though the association of sex steroid hormones with breast cancer risk is well studied and their association with gonadotrophins in menopausal transition has been well described, interaction of sex steroids and gonadotrophins in breast cancer is not known. Moreover, there are not any studies investigating whether combination of hormones could improve risk prediction further investigating their possible

synergistic effects in breast carcinogenesis. The only study that has reported data on combinational effect of endogenous oestrogens and androgens on breast cancer risk was by Adly *et al*, showing a higher increased risk for women having oestrone sulphate and androstenediol in top quintiles compared to each single hormone [201]. It needs to be pointed though that the samples used for the purposes of this study were taken at the point of diagnosis and not years before diagnosis as the samples used in this study. Therefore, we hypothesised that by investigating different combinations of sex steroid hormones, gonadotrophins and SB of sex steroid receptors could prove to have a better breast cancer risk prediction power in comparison to each individual measurement and provide information on their synergistic effect in breast carcinogenesis.

Using the UKCTOCS biobank we were able to explore all the above issues to better understand breast carcinogenesis. Women recruited to the trial between 2001-2005 provided blood samples for secondary studies and continue to be followed up by cancer registration and self-reporting [78, 335]. A nested case control study was undertaken using serum samples donated between 6 months and 5 years before diagnosis by women who developed breast cancer after joining the trial and healthy women who had not developed the disease examining: **1)** Five sex steroid hormones (oestradiol, oestrone, androstenedione, testosterone, DHEAS), free oestradiol and free testosterone (calculated by the mass action law), two gonadotrophins (LH and FSH) and SHBG in association with ER-positive breast cancer risk. Since all earlier studies have only explored levels of endogenous hormones with regard to breast cancer risk and with the relatively new discovery of very sensitive bioactivity assays for steroid hormones being able to detect very low hormone levels, we investigated **2)** SB of ER-α and -β and

AR in breast cancer and examined whether they are associated and predict the disease. **3)**. Since previous studies have looked at the effect of each individual hormone in breast cancer risk, joint associations of sex steroids, gonadotrophins and steroid receptor bioactivity were examined hypothesising that they may have better risk prediction and further examined their synergistic effect in breast cancer. Moreover, **4)** association of hormones and serum bioactivity of the receptors in relation to time of breast cancer diagnosis was also investigated.

4.2 Methods and Materials

4.2.1 Eligible cases and samples

Breast Cancer Cases were women identified in chapter 3, who fulfilled the eligibility criteria stated below:

(1) ER-positive invasive breast cancer diagnosis

(2) not having HRT treatment at recruitment and

(3) having a serum sample given at least 6 months up to 5 years prior to diagnosis following randomisation into the trial

Controls were women who had:

(1) no history of breast cancer and any other cancer

(2) had a serum sample collected on the same day and in the same clinic as the cases

Two controls were selected for each case and were age matched to breast cancer cases.

Blood samples were collected in Greiner gel tubes (Cat no: 455071) at the centres and couriered overnight to the central UKCTOCS laboratory. The samples were centrifuged at 2000 g for 10 minutes and the serum was removed from the cells within 56 hours of sample collection. A novel semi automated system aliquoted serum in 500 micro liter straws which were then heat sealed bar coded and stored in special containers in liquid nitrogen tanks. Two straws were retrieved, one for the measurement of hormonal levels and one for the bioactivity assays. The samples were only thawed before use.

4.2.2 Collection of epidemiological factors

As mentioned in methods and materials in chapter 3, there were two large-scale questionnaire surveys covering demographics, health behaviour, medical history and epidemiological factors were conducted during UKCTOCS trial. One was based at the time of recruitment and one after 3.5 years of participation into the trial. From the questionnaires the following (potential) breast cancer risk factors were obtained: ethnicity, height input, weight input, height, BMI calculated as weight (in kilograms) divided by height (in metres) squared, age at first period, age at menopause, skirt size difference (increase/decrease), ovarian cancer family history, breast cancer family history, HRT use, hysterectomy, pill use, pregnancies less than 6 months, pregnancies more than 6 months, sterilisation, infertility.

4.2.3 Sex steroid hormonal levels using immunoassay systems

Enzyme-Linked ImmunoSorbent Assay (ELISA) is the most popular immunological assay because of its versatility, sensitivity, specificity and ease of automation. It is a biochemical technique used to detect the presence of an antibody or an antigen in a sample. In simple terms, in ELISA an unknown amount of antigen is affixed to a surface, and then a specific antibody is washed over the surface so that it can bind to the antigen. This antibody is linked to an enzyme, and in the final step a substance is added that the enzyme can convert to some detectable signal. For the purposes of this study two different types of ELISA assays were used; the sandwich and competitive assay. A standard curve with known concentrations of the antigen of interest is plotted in order to determine the unknown antigen in experimental samples.

Kits for SHBG, LH and FSH (electrochemiluminescence sandwich immunoassays), oestradiol, testosterone, DHEAS and progesterone (electrochemiluminescence competitive immunoassays) were obtained from Roche Diagnostics and the samples assayed on an Elecsys 2010 analyzer (Roche Diagnostics GmbH, Mannheim, Germany). Androstenedione was analysed by competitive chemiluminescent immunoassay on DPC IMMULITE 2500 analyzer (SIEMENS Medical Solutions Diagnostics, Germany). For oestrone, ELISA kits (solid phase competitive enzyme immunoassay on microtitre plates) were obtained from DRG (DRG, Instruments GmbH, Germany). Information regarding the kits and how the assays were performed in detail are provided in Appendix IV, specifications of the ELISA kits are provided in Table 4-1. The samples were analysed blind in randomly mixed batches of cases and controls using a single lot number of reagent and calibrator. All measurements were done by me.

Briefly the principles of the different assays used were: Competitive assay - Elecsys 2010 analyser, samples were incubated with biotinylated monoclonal specific antibody and a monoclonal specific antibody labelled with ruthenium. The binding sites of the labelled antibody became occupied partially by the sample analyte (depending on its concentration) and partly by the ruthenium-labelled hapten forming the respective immunocomplexes. Sandwich assay - the samples were incubated with biotinylated monoclonal specific antibody and a monoclonal specific antibody labelled with ruthenium forming a sandwich complex. After the addition of the streptavidin coated microparticles the complex became bound to the solid phase. The reaction mixture was aspirated into the measuring cell where the microparticles were magnetically captured into the surface of the

electrode. For DPC IMMULITE 2500 analyser, the antigen in the sample competed with a fixed amount of alkaline phosphatase-conjugated label to bind with a polyclonal rabbit antibody coated solid phase (polystyrene bead).

Table 4-1: Specifications of ELISA kits.

Hormone	Sample (µl)	Measuring Range	Intra-Assay Variation (%)	Inter-Assay Variation (%)
Oestradiol	35	18.4-15,781 pmol/L	1.6-5.7	2.3-6.2
Oestrone	25	15-2000pg/ml	4.5-9.3	7.4-12.9
Androstenedione	25	1-35 nmol/L	3.5-11.3	4.4-13.2
Testosterone	50	0.0695-52 nmol/L	0.9-4.6	1.6-7.4
DHEAS	15	0.003-27 ulmol/L	0.8-1.8	1.9-5.2
SHBG	10	0.350-200 nmol/L	2.1-2.7	2.6-5.6
Progesterone	30	0.095-191 nmol/L	1.5-2.7	3.7-5.4
LH	20	0.100-200 mIU/mL	1.7-2.8	2.4-4.7
FSH	40	0.100-200 mIU/mL	1.4-2.0	2.9-5.1

DHEAS=dehydroepiandrosterone sulphate; ELISA=enzyme-linked immunosorbent assay; FSH=follicular stimulating hormone; LH=luteinising hormone; SHBG=sex hormone-binding globulin

For all assays used the unbound substances were then removed with Procell (Elecsys 2010 analyser) or L2KPM (DPC IMMULITE 2500 analyser). Application of a voltage to the electrode then induced chemiluminescent emission which was measured by photomultiplier. The results were determined via a calibration curve which was instrument specifically generated by 2-point calibration and a master curve that was provided via the reagent barcode. The calibration procedure was performed before running the samples in the analysers. For the method the stored master curve adjusted by running the low and high adjusters was carried each in replicates.

Competitive assay was used to measure oestrone levels. The antigen of the sample competed with oestrone horseradish peroxidase conjugate for binding to

the coated antibody. The amount of bound peroxidase conjugate was reverse proportional to the concentration of the oestrone in the sample. Therefore, after adding the substrate solution the intensity of the colour developed was reverse proportional to the concentration of oestrone in the sample. The samples, controls and standards (for standard curve) were run in duplicates. In order to construct the standard curve the mean absorbance obtained from each standard (on the vertical (Y)) was plotted against its concentration (concentrations 0-15-50-200-800-2000 pg/ml) (on the horizontal (X)). The best fitted curve was obtained by using a 4 parameter logistics curve fit (Excelstat). The concentrations of the samples were read directly from this standard curve.

Quality controls were run on each day for the samples that were done in the analysers and they were included in each plate for the samples that were done manually. Details regarding the quality controls are provided in Appendix V. For the samples run in Elecsys 2010 analyser, PreciControl Universal PC1 and PC2 were used, for the samples run in DPC IMMULITE 2500 analyser, CO6 was used and for the samples run manually to measure oestrone levels controls were provided within the kit.

4.2.4 Calculation of free oestradiol and testosterone

For the calculation of free oestradiol (fE_2) and free testorenone (fT) the equation based on mass of action law by Vermeulen [336] was used. The equation relies on the assumption that the concentration of fE_2 and fT in blood is determined mainly by the interaction between SHBG and albumin, and that other hormones present in the blood do not influence this equilibrium much.

Equations:

$$[fE_2] = \frac{([E_2] - (N1 \times [fE_2]))}{(K_s E_2 \{[C_{SHBG}] - [E_2] + N_2[fE_2]\})}$$

$$[fT] = \frac{([T] - (N2 \times [fT]))}{(K_s T \{[C_{SHBG}] - [T] + N_1[fT]\})}$$

where $[E_2]$ and [T] are total oestradiol and testosterone concentrations; K_SE_2 and K_ST are the affinity constants for SHBG for E_2 and T; $N1 = K_aE_2C_a + 1$ and $N2 = K_aTC_a + 1$, where C_a is the albumin concentration and K_aE_2 and K_aT are the affinity constants of albumin for E_2 and T.

$K_SE_2 = 3.14 \times 10^8$ liters/mol $K_ST = 1 \times 10^9$ liters/mol

$K_aE_2 = 4.21 \times 10^4$ liters/mol $K_aT = 4.06 \times 10^4$ liters/mol

$C_a = 6.5 \times 10^{-4}$ mol/litre

4.2.5 Sex steroid hormonal receptor bioactivity assay

The test used to measure sex steroidal hormonal receptor bioactivity is a yeast based reporter gene assay which not only determines whether a chemical binds to the receptor, but also if oestrogen or androgen-dependent gene expression is stimulated. The recombinant yeast was provided from our collaborators from University of Bonn, Germany where all the experimental work was carried out after being trained. The group run by Professor Hella Lichtenberg-Fraté has published results based on the assay [337, 338]. The assay utilises the yeast *Saccharomyces cerevisiae*, an eukaryotic organism, as the biological component since it has been proven to be a good model for studying more complex eukaryotic processes, such as steroid receptor function. Yeast is an attractive

and widely used model because the cellular structure is that of eukaryotes, like mammals. It exhibits a eukaryotic architecture with internal organelles and similar chromosome structure and DNA repair and metabolic processes. Therefore, by using such an assay it allows the combination of a eukaryotic test system with the advantages connected to prokaryotic systems like the short incubation time, reproducible growth rates, simple optical read outs and the ability to use well defined culture conditions.

The test was developed by preparing different strains of genetically modified yeasts integrating the DNA sequence for the human ER-α or ER-β or AR-b into the main chromosome of *Saccharomyces cerevisiae*. The recombinant yeast cells contain HRE where the ligand binds and a plasmid that possesses the *Aquorea victoria* green fluorescence protein (GFP) as reporter gene. This method is applicable to complex samples (blood serum) which are soluble under the conditions of the test. The endpoint is the determination of fluorescence development. Upon exposure of the genetically modified *Saccharomyces cerevisiae* cells the production of GFP which, upon excitation by 485 nm emits green fluorescence whose emission at 535 nm can be detected using photodetectors. Results are obtained in arbitrary fluorescence units versus increasing 17β-E2 concentrations for the ER-α and ER-β bioactivity assay and DHT for the AR-b bioactivity assay (calibration curve).

Briefly, the genetically modified yeast cells are incubated in a defined test medium with the reference substance 17β-E2 or DHT and different test samples. At the end of the incubation period the developed green fluorescence is determined and corrected for cell density, optical density (OD) of the cell

suspension and blanks. The measurements were performed in microplate reader (TECAN). The cell growth was determined by measuring the light absorption at 600 nm and GFP-fluorescence was determined by measuring GFP at 535 nm, specific OD and fluorescence at t = 0 and t = 16,5 h for ER-α and ER- β and t = 24h for AR-b was measured in each of the 96-wells. Tests were considered as valid if the turbidity of the negative control culture increased five times during the incubation period. The control culture should expose no fluorescence development. The bioactivity was determined by comparison of the fluorescence development in test cultures versus the 17β-E2 or DHT calibration curve. The dose-response curves of the reference values were fitted using the Hill equation fit and the R function (The R Foundation for Statistical Computing, http://www.r-project.org/) [337, 338]. Analysis was performed blind and cases and controls were randomly mixed. Tests were carried out with two replicates at a time on two different days (thus four readings in total. No temporal effects were detected. By the nature of the assay, minor daily performance differences may occur, but were accounted by including a daily reference curve, comprising 10 different concentrations. Order effects were not detected since, as mentioned above, all samples were randomised before numerical coding.

Media preparation

Amino Acid-Drop-out-Mix: L-Arginine (100 mg), L-Methionine (100 mg), L-Tyrosine (100 mg), L-Lysine (150 mg), L-Valine (300 mg), L-Threonine (500 mg), L-Serine (500 mg), L-Phenylalanine (250 mg), L-Asparagin (100 mg), L-Glutamic acid (100 mg), Adenine (250 mg), L-Histidine (100 mg). All the components were added in a glass container and mixed thoroughly.

5x-concentrated liquid nutrient medium-Yeast Nitrogen Base (YNB) medium with 0.5% glucose, pH 6.4: Components for 1000 ml medium were mixed. Yeast Nitrogen Base (YNB) (DIFCO) w/o amino acids and w/o ammonium sulfate (8.5 g), Amino Acid Drop-Out-Mix (2.6 g), Ammonium-nitrate (25 g), Citrate-buffer (50 mM final concentration, 52.5 g). Sterilised water was added to 850 ml. The pH was adjusted to 6.4 by adding 25-30 g sodium hydroxide pellets and subsequently 5 M sodium hydroxide-solution. Sterilised water was added to 937.5 ml and the medium was autoclaved (20 minutes at 121°C). The addition of autoclaved stock solution of 62.5 ml 40% glucose (in sterilised water) was conducted under a clean-bench.

Charcoal stripped serum preparation

The protocol was taken from Miller *et al*, 1999 [339]. A mixture of 0.5% charcoal and 0.05% dextran in 50 mM HEPES (buffering agent), ph 8.0 was smoothly agitated for 30 minutes at 37 °C. The slurry was centrifuged at 3000 rpm, 4 °C, for 30 minutes and the supernatant was removed and replaced with foetal bovine serum that had been heat-inactivated for 30 minutes at 56 °C. The mixture was then smoothly agitated for 3 h at 37 °C and finally centrifuged at 3000 rpm, 4 °C for 60 minutes. After centrifugation the serum was carefully pipetted away from the formed charcoal pellet and filter sterilised. The serum was stored at -20 °C until used.

Preparation of 17β-E2 and DHT stock solutions

Water soluble 17β-E2 (Sigma E4389) was dissolved in stripped serum to a final concentration of 1mg/ml. Solution was stored at -20 °C. From this solution a dilution series was prepared with 1:50 steps. From the dilution series 17β-E2

stock solutions were prepared by adding stripped serum to give final concentrations in the test of [17βE2] = 0, 1, 3, 10, 30, 100, 300, 1000, 3000, 10000 pg/ml. The 5 x 17β-E2 stock solutions were directly in a volume of 20µl (according 20 %) to the test culture of overall 100µl. Dilution series and 5x17β-E2 stock solutions were stored at 4 °C for 4 weeks. The same procedure was followed for DHT (Sigma A8380).

Yeast starter and pre-culture cultivation

Vials containing the recombinant yeast strains with sterile glycerol were obtained from our collaborators which were kept frozen at -80 °C. From all yeast strains, 20 µl of the glycerol stock were spread out on YNB nutrient medium agar plates (prepared by our collaborators). The lid of the agar plate was sealed with parafilm and incubated at 30 ° C until yeast colonies were observed.

For the liquid pre-cultures, cell material from one selected single colony was taken with a sterile toothpick and inoculated in 20 ml YNB-medium in a sterile 100ml flask with cellulose stopper or metal tight-lock cap. The flask was incubated on a rotary shaker at 30 °C and 220 rpm until the cells entered the stationary phase (overnight, 16 h). The cell density was above 4.5 x 10E7 cell/ml and the visual inspection (microscope, 600x) of the cells resulted in a low percentage (<20%) of budding cells. Such a yeast culture was stored at 4 °C and was used as pre-culture for subsequent tests for up to 5 days maximum.

Conduction of the assay

The prepared liquid pre-culture was centrifuged (3000 rpm for 3 min) and the supernatant was removed and the cells were resuspended in 1 ml sterile water

and centrifuged again. The supernatant was removed again and the cells were resuspended in 1 ml 5x YNB-medium. The cell density was determined by means of OD 600 measured in a photometer. The final cell density was adjusted with the 5 x YNB-medium of 4 x 10E7 cells/ml. The serum samples were homogenised by vigorous shaking immediate before usage. 80 µl of the cell suspension were pipetted in each (culture) well. 20 µl of testing serum or internal standard curve serum to the (culture) wells were added. Negative controls were added in the plate consisting of 80 µl cell suspension of the cells and 20 µl of water. The OD and fluorescence at time zero (t = 0) was measured in the reader and the plate was sealed with lid and with parafilm. Finally, the plate was placed on a rotary shaker with 950 rpm for 16, 5 hours (ER-α and ER-β) or 20 hours (AR-b). Incubation temperature was 30 °C. Measurement of the OD at 600 nm and of the GFP-specific fluorescence emission at 535 nm of all wells was conducted after the incubation period.

Data evaluation

After 16.5 hours or 20 hours incubation depending on the assay, the obtained end point fluorescence (FL) values (corrected for blanks) were divided by growth determined as OD (corrected for blanks) for each replica well to normalise fluorescence for cell number (FL/OD).

To increase the reproducibility of results, the FL/OD values obtained for a test compound at a given concentration were expressed as fractional values of the maximal response of a saturating concentration of the reference compound E2 or DHT (internal standard curve).

The top and bottom values were obtained by Hill equation fit using the R function:

$$y(x) = \text{bottom} + \frac{\text{top - bottom}}{1 + 10^{(\text{LEC50-x})\cdot\text{hillslope}}}$$

with y(x) = FL/OD at the actual compound concentration, x = the decadical logarithm of compound concentration, LEC50 = decadical logarithm of EC50, top = fitted maximal FL/OD at saturating concentrations, bottom = fitted maximal FL/OD of negative control and hill_slope as the hill steepness parameter to the FL/OD values for each E2 or DHT concentration.

4.2.6 Statistics

Mean and median levels of sex steroid hormones, ER-α and ER-β and AR SB were calculated for all breast cancer samples and for controls. Differences in the medians between groups were tested for statistical significance using the Kruskal–Wallis test. Correlations between sex steroid hormones and, ER-α and ER-β and AR SB among cases and controls were assessed by the Spearman's rank correlation coefficient. Statistical analysis was carried out using a computer assisted program-SPSS version 12.0.1, Chicago, IL. The associations between hormones/SB of sex steroid receptors and risk of breast cancer was determined by logistic regression to estimate OR and compute 95% confidence intervals (CI). Subjects were classified according to quintiles of the respective marker among controls. Cut-off points of the top quintiles for the different hormones and SBs are provided in table 4-2. Hormones and sex steroid receptor SB levels were entered and controlled in regression models to estimate their independent and combined associations with breast cancer risk. All regression analyses were adjusted for

age. Therefore, we present ORs not adjusted (only age adjustment) and adjusted for other serum hormones/SB to aid in understanding their relationship to one another.

Table 4-2: Cut-off points of the top quintiles for sex steroids, sex hormone-binding globulin, gonadotrophins and serum bioactivity of sex steroid hormone receptors.

Hormone	Cut-off point
Oestradiol (pg/ml)	22.66
Free Oestradiol (pmol/l)	1.191
Oestrone (pg/ml)	115.532
Androstenedione (nmol/l)	4.614
Testosterone (nmol/l)	0.382
Free Testosterone (ng/dl)	0.164
DHEAS (ug/dl)	162.04
SHBG (nmol/l)	80.134
Progesterone (ng/ml)	0.387
LH (mIU/ml)	42.978
FSH (mIU/ml)	98.32
ER-α (pg/ml)	104.359
ER-β (pg/ml)	98.955
AR (ng/ml)	2.867

AR=androgen receptor; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; SHBG=sex hormone-binding globulin

Further analysis was carried out to investigate the effect of paired hormones/SB on breast cancer risk prediction. Subjects were identified with different pairs of hormone/SB levels in the highest quintile and compared to those that did not have at least one variable in the highest quintile (the predictor variable construction is presented in Formula (1). For some of the investigated pairs not enough points were found in the highest quintiles for both hormones/SB. Therefore, for these pairs instead of the OR and CIs values the description ‘‘not

enough points'' is given. All regression analyses were adjusted for age. Further investigation was undertaken by adjusting for each individual hormone/SB. Therefore, the data is presented as ORs unadjusted (only age adjustment) and adjusted for other serum hormones/SB. To validate the results on the best pairs identified through the above described analysis 1000 experiments were run, where 10% of the data was removed from cases and 10% of controls. The quintiles were re-calculated and ORs were re-evaluated. Mean, median and variance were calculated to examine the distribution of the data.

Next the synergistic effect of the different pairs (hormones/SB) was investigated. The ratio of observed versus expected was computed quantifying the hidden synergistic effect of hormones/SB pairs. Initially, the expected value of the OR for the different pairs was calculated based on the value of the single observed OR and the corresponding regression coefficients. This was computed using the algorithm to construct the predictor variable (Formula 1) and under the assumption that hormones/SB are independent. The expression that defined the expected OR as a function of the single OR is demonstrated in Equation 1. To compute the CIs of the expected ORs and the CIs of ratio of observed versus expected ORs, Monte Carlo stimulation was used. For analysis the R Foundation for Statistical Computing program was used. Description of the models created by Professor Alexey Zaikin is given below:

Encoding the predictor variable:

Pair of hormones/SB

For pair of hormones the predictor variable Xij, i=1 N the predictor variable Xi is constructed as follows:

$$X_{ij} = \begin{cases} 1, & if \quad H_i \; and \; H_j \quad are \quad in \quad 5th \quad Quintile \\ 0, & if \quad otherwise \end{cases}$$ Formula (1)

Quintile values are chosen on the base of the control set only.

Calculation of the OR for the pair hormones/SB:

Correspondingly, for the predictor variable Xij describing the joint action of two hormones/SB i and j OR was calculated from the regression of the logistic $log\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 X_{ij}.$ as $OR = exp(\beta_1)$, or from $log\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 X_i + \beta_2 X_j + \beta_3 A.$ as $OR = exp(\beta_1)$ If adjustment was carried out by another hormone/SB predictor Xj and age predictor variable A.

Calculation of the expected OR:

Suppose that for the two hormones the following two probability tables.

To calculate the expected OR when the two ORs when the single hormones/SB OR1 and OR2 and the corresponding interceptors R1 and R2, defined as Ri=exp (βi0) are known and since the coefficients of the logistic regression are known, the probability to be in Case category for any value of the predictor variable can be found.

For the two different values of the predictor variable:

$$p1^1 = \frac{R_1}{1+R_1} \quad p0^1 = \frac{R_1 OR_1}{1+R_1 OR_1}.$$

Using this probabilities and solving a system of equations

$$p1^1 = \frac{p_{11}^1}{p_{11}^1 + p_{10}^1}, \quad p0^1 = \frac{p_{01}^1}{p_{01}^1 + p_{00}^1}, \quad N' = p_{11}^1 + p_{01}^1 = p_{00}^1 + p_{10}^1 + p_{01}^1 + p_{11}^1$$

Where N'=Ncases/N, the table probabilities can be found

$$p_{01}^1 = -\frac{-R_1 OR_1 + N' + R_1 OR_1 N}{-1 + OR_1},$$
$$p_{00}^1 = -\frac{-R_1 OR_1 + N' + R_1 OR_1 N}{R_1(-1 + OR_1)},$$
$$p_{10}^1 = -\frac{N + NR_1 - R_1}{R_1(OR_1 - 1)},$$
$$p_{11}^1 = -\frac{(N + NR_1 - R_1)OR_1}{(OR_1 - 1)}.$$

Using these table probabilities the probability of the predictor variable to be equal to 1 among cases was calculated.

$$pc^1 = \frac{p_{11}^1}{p_{11}^1 + p_{01}^1} = \frac{(N' + N'R_1 - R_1)OR_1}{OR_1 - 1}$$

And among controls $pk^1 = \frac{p_{10}^1}{p_{10}^1 + p_{00}^1} = -\frac{(N' + N'R_1 - R_1)}{R_1(1 - N' - OR_1 + N'OR_1)}.$

Following the same method the probabilities for the second hormones/SB were found.

$$pc^2 = \frac{p_{11}^2}{p_{11}^2 + p_{01}^2} = \frac{(N' + N'R_2 - R_2)OR_2}{OR_2 - 1}$$

And among controls $pk^2 = \frac{p_{10}^2}{p_{10}^2 + p_{00}^2} = -\frac{(N' + N'R_2 - R_2)}{R_2(1 - N' - OR_2 + N'OR_2)}.$

Using the same methods of the introduction of the predictor variable for the hormone/SB pairs, the probabilities of the joint predictor variable was estimated.

$pc^3 = pc^1 pc^2 \quad pk^3 = pk^1 pk^2.$

Knowing these probabilities a system of equations was solved

$$pc^3 = \frac{p_{11}^3}{p_{11}^3 + p_{01}^3}, \quad pk^3 = \frac{p_{10}^3}{p_{00}^3 + p_{10}^3}, \quad N = p_{11}^3 + p_{01}^3, \quad 1 = p_{00}^3 + p_{10}^3 + p_{01}^3 + p_{11}^3$$

And the probabilities for the hormone/SB pair were found: $p_{00}^3, p_{10}^3, p_{01}^3, p_{11}^3.$

Using these probabilities, the following calculations were made to get:

$OR_3 = \frac{p_{11}^3 p_{00}^3}{p_{10}^3 p_{01}^3}$ and to obtain $OR_3 = OR_1 OR_2 \frac{T}{B}$, where;

$$T = -(-OR_2 N'^2 R_1 R_2 - OR_1 N^2 R_1 R_2 - N'^2 R_2 - N'^2 R_1 - N'^2 + OR_1 OR_2 N'^2 R_1 R_2 + N' R_1 - 2OR_1 OR_2 R_1 N' R_2 + N' R_2 + 2OR_2 R_1 N' R_2 + 2OR_1 R_1 N' R_2 - OR_2 R_1 R_2 + OR_1 OR_2 R_1 R_2 - OR_1 R_1 R_2) \quad \text{Equation (1)}$$

And

$$B = (OR_1 OR_2 R_1 R_2 - OR_1 OR_2 N' R_1 - 2OR_1 OR_2 R_1 N' R_2 - OR_1 OR_2 N' R_2 + OR_1 OR_2 N'^2 R_2 + OR_1 OR_2 N'^2 R_1 + OR_1 OR_2 N'^2 R_1 R_2 - N'^2 + N'^2 OR_2 + N'^2 OR_1)$$

To summarise, joint OR for two hormones/SB as a function of their single OR (OR_1, OR_2) and their interceptor coefficients (R_1, R_2) and proportion of cases $N' = N_{cases} / N$ was found. For analysis the R Foundation for Statistical Computing program was used.

4.3 Results

4.3.1 Clinicopathological characteristics of the eligible cases

Clinicopathological characteristics of the breast cancer cases are provided in Table 4-3. Cases included 200 women with invasive breast carcinoma. Most of the tumours were ductal (81%), 48% were early stage (stage 1) at diagnosis and only 5% were advanced, with 27.5% being un-staged. Regarding hormone status, all of the cases were ER-positive, 50% were PR-positive and 39.5% were HER2-negative, with the majority of the cases having an unknown HER2 status as it was not routinely performed in the hospitals where women were treated.

Table 4-3: Clinicopathological details of cases.

Histology Classification	**No**	**%**
IDC	162	81
ILC	25	12.5
ITC	1	0.5
Other	12	6
Stage (TNM)		
1	96	48
2	39	19.5
3	10	5
Unknown	55	27.5
Grade		
I	32	16
II	111	55.5
III	53	26.5
Unknown	4	2
Oestrogen Receptor		
ER positive	200	100
Progesterone Receptor		
PR negative	32	16
PR positive	100	50
Unknown	68	34
HER2		
HER2 negative	79	39.5
HER2 positive	16	8
Unknown	105	52.5
Nodal Status		
Positive	50	25
Negative	150	75

HER2=human epidermal growth factor receptor 2; IDC=invasive ductal carcinoma; ILC=invasive lobular carcinoma; ITC=invasive tubular carcinoma

4.3.2 Epidemiological risk factor profile of the study women

The median age of the 200 women with breast cancer (cases) was 61.33 (inter-quartile range IQR 11.32) and 62.33, (IQR 9.57) in the 400 healthy women (matched controls). None of the traditional risk factors (family history, age at menarche, menopause, number of pregnancies, contraceptive pill use, hysterectomy, infertility, BMI, height) were significantly different between cases and controls except for fallopian tube ligation (OR for breast cancer 0.57; 95% confidence interval (CI): 0.35-0.94; p=0.029) (Table 4-4).

Table 4-4: Traditional risk factors in cases and controls. (numbers of cases and controls do not always add up to totals due to missing values in some participants; cases N=200 and controls N=400)

Characteristics	Category	Case N	%	Control N	%	Total N	OR*	L95%CI	U95%CI	p-value
Ethnicity	**Non-white**	5	2.5%	11	2.7%	16	1.00			
	White	194	97.5%	392	97.3%	586	1.09	0.37	3.18	0.876
Breast cancer family history	**No**	148	74.4%	291	76.6%	439	1.00			
	Yes	51	25.6%	89	23.4%	140	1.13	0.76	1.68	0.556
Ovarian cancer family history	**No**	189	95.0%	363	95.5%	552	1.00			
	Yes	10	5.0%	17	4.5%	27	1.13	0.51	2.52	0.765
Age at menopause	**<50**	85	42.7%	187	46.4%	272	1.00			
	50+	114	57.3%	216	53.6%	330	1.16	0.82	1.64	0.392
Age 1st period	**<12**	47	23.6%	79	19.8%	126	1.00			
	12+	152	76.4%	319	80.2%	471	0.80	0.53	1.21	0.288
Pregnancies <6 months	**None**	31	15.6%	50	12.5%	81	1.00			
	1+	168	84.4%	349	87.5%	517	0.78	0.48	1.26	0.306
Pregnancies >6 months	**None**	146	73.4%	273	69.1%	419	1.00			
	1+	53	26.6%	122	30.9%	175	0.81	0.56	1.19	0.284
HRT use	**No**									
	Yes	200	100%	400	100%					
Pill use	**No**	98	49.2%	180	44.7%	278	1.00			
	Yes	101	50.8%	223	55.3%	324	0.83	0.59	1.17	0.289
Hysterectomy	**No**	162	81.4%	322	79.9%	484	1.00			
	Yes	37	18.6%	81	20.1%	118	0.91	0.59	1.40	0.661
Infertility	**No**	195	98.0%	393	97.5%	588	1.00			
	Yes	4	2.0%	10	2.5%	14	0.81	0.25	2.60	0.719
Sterilization	**No**	176	88.4%	328	81.4%	504	1.00			
	Yes	23	11.6%	75	18.6%	98	**0.57**	0.35	0.94	**0.029**
BMI	**<24.0**	54	27.1%	132	33.2%	186	1.00			
	24.0+	145	72.9%	266	66.8%	411	1.33	0.92	1.94	0.134
Height	**1.6m**	89	44.7%	191	48.0%	280	1.00			
	1.6m+	110	55.3%	207	52.0%	317	1.14	0.81	1.61	0.451

BMI= body mass index; HRT=hormone replacement therapy; OR=odds ratio

4.3.3 Association of hormones and serum bioactivity with breast cancer

Amongst the nine hormones analysed difference between cases and controls was observed for serum androstenedione, testosterone and free testosterone levels. (Table 4-5 and Table 4-6). Women were then stratified into groups based on whether the sample was obtained 6 months to ≤2 or >2 to 5 years prior to breast cancer diagnosis. For those women who had given samples ≤2 years before diagnosis, SHBG and serum free testosterone showed significant differences between cases and controls (Table 4-5 and Table 4-6). For those women who had given a sample >2 years before diagnosis, androstenedione, testosterone and free testosterone showed significant differences between cases and controls (Table 4-5 and Table 4-6). The other hormones did not show statistically significant differences between cases and controls (Table 4-5 and Table 4-6).

A significant difference for both ER-α and ER-β SB was shown between cases and controls for those samples that were taken >2 years before diagnosis but not when all cases were investigated or for those women that gave samples ≤2 years before breast cancer diagnosis. AR SB did not show any statistically significant difference between cases and controls (Table 4-7).

Table 4-5: Oestrogens and androgens levels in serum samples from 200 cases and 400 controls. (The numbers do not always add up due to some missing values)

Hormone	Case-Control Status	Number	Mean	Median	Standard Deviation	p-value*
Oestradiol (pg/mL)	Controls	379	18.44	16.03	13.81	
	All Cases	194	18.57	16.51	10.59	0.47
	≤2 years	93	17.93	16.24	11.19	0.93
	>2 years	100	19.20	16.87	9.96	0.22
Free oestradiol (pmol/l)	Controls	362	0.91	0.79	0.62	
	All Cases	193	0.98	0.84	0.51	0.07
	≤2 years	93	0.93	0.84	0.44	0.17
	>2 years	100	1.00	0.84	0.57	0.17
Oestrone (pg/ml)	Controls	384	99.74	80.93	80.63	
	All Cases	198	112.42	81.79	125.44	0.11
	≤2 years	95	116.56	83.14	132.79	0.09
	>2 years	103	108.72	81.16	118.22	0.46
Androstenedione (nmol/L)	Controls	386	3.38	3.13	1.76	
	All Cases	195	4.07	3.59	2.30	**0.01**
	≤2 years	95	3.95	3.37	2.27	0.19
	>2 years	100	4.19	3.71	2.33	**0.00**
Testosterone (nmol/L)	Controls	382	0.28	0.25	0.16	
	All Cases	193	0.35	0.28	0.28	**0.04**
	≤2 years	94	0.31	0.27	0.18	0.08
	>2 years	99	0.38	0.29	0.34	**0.01**
Free Testosterone (ng/dl)	Controls	365	0.12	0.09	0.20	
	All Cases	193	0.15	0.11	0.13	**0.00**
	≤2 years	93	0.13	0.11	0.10	**0.03**
	>2 years	100	0.16	0.14	0.16	**0.03**
DHEAS (ug/dl)	Controls	385	111.83	100.60	61.15	
	All Cases	195	118.87	97.95	72.92	0.58
	≤2 years	95	121.30	103.00	70.78	0.25
	>2 years	100	116.50	93.90	75.31	0.95
SHBG (nmol/L)	Control	385	56.62	53.25	26.09	
	Case	195	50.67	47.55	20.98	0.12
	≤2 years	95	48.54	47.40	20.49	**0.02**
	>2 years	100	52.79	47.71	21.36	0.50
Progesterone (ng/ml)	Control	382	0.29	0.25	0.26	
	Case	139	0.38	0.24	1.07	0.60
	≤2 years	195	0.42	0.24	1.29	0.32
	>2 years	95	0.29	0.24	0.21	0.63

*Kruskal-Wallis for difference in median value among cases and controls.

DHEAS=dehydroepiandrosterone sulphate; SHBG=sex hormone-binding globulin

Table 4-6: Gonadotrophin levels in serum samples from 200 cases and 400 controls.

(The numbers do not always add up due to some missing values)

Hormone	Case-Control Status	Number	Mean	Median	Standard Deviation	p-value*
LH (mIU/ml)	Controls	387	33.11	30.84	13.72	
	All Cases	195	31.64	30.53	11.16	0.39
	≤2 years	100	32.07	31.37	10.58	0.95
	>2 years	95	30.50	29.12	11.10	0.10
FSH (mIU/ml)	Controls	387	78.18	76.95	28.52	
	All Cases	195	75.58	74.27	27.58	0.64
	≤2 years	100	75.49	74.43	22.34	0.69
	>2 years	95	75.66	73.94	32.10	0.48

*Kruskal-Wallis for difference in median value among cases and controls.

FSH=follicular stimulating hormone; LH=luteinising hormone

Table 4-7: Serum bioactivity of oestrogen receptor-α and -β and androgen receptor in serum samples from 200 cases and 400 controls. (The numbers do not always add up due to some missing values)

SB	Case-Control Status	Number	Mean	Median	Standard Deviation	p-value*
ER-α (pg/ml)	Controls	390	70.74	62.09	60.45	
	All Cases	198	80.24	64.17	68.54	0.30
	≤2 years	95	74.86	57.60	69.81	0.78
	>2 years	103	85.60	74.85	67.19	**0.05**
ER-β (pg/ml)	Controls	391	59.95	43.87	67.63	
	All Cases	198	71.69	48.22	83.79	0.41
	≤2 years	95	61.10	37.56	85.81	0.26
	>2 years	103	82.26	59.64	80.79	**0.01**
AR (ng/ml)	Controls	391	2.33	2.32	1.01	
	All Cases	197	2.36	2.29	0.86	0.97
	≤2 years	94	2.28	2.26	0.85	0.20
	>2 years	103	2.44	2.38	0.88	0.19

*Kruskal-Wallis for difference in median value among cases and controls.

AR=androgen receptor; ER=oestrogen receptor; SB=serum bioactivity

4.3.4 Association of hormones and serum bioactivity with breast cancer risk

Further analysis was carried out of the data based on top-bottom classification, using top quintile versus bottom 4 quintiles, according to top-bottom classification among controls. When all cases were used for the analysis significant association with serum androstenedione (≥4.614 nmol/L), testosterone (≥0.382 nmol/L), free testosterone (≥ 0.164 ng/dl) and SHBG (≥ 80.134 nmol/L) was observed in breast cancer samples. Women who had serum levels in the top quintile of androstenedione, testosterone and free testosterone had 1.854 (95% CI: 1.240-2761), 2.238 (95% CI: 1.512-3.317), 1.637 (95% CI: 1.090-2.543) fold breast cancer risk. In order to test whether these hormones were independently associated with breast cancer risk, a logistic regression analysis was performed adjusting for the other hormones and SB of steroid receptors. The association of testosterone and SHBG with breast cancer remained significant throughout all adjustments. However, the association of androstenedione with breast cancer risk was not statistically significant after adjustment for testosterone and for free testosterone after adjustment for oestradiol and androstenedione, respectively. DHEAS was only significantly associated with breast cancer risk after adjustment for testosterone. Oestrogens – oestradiol, free oestradiol and oestrone did not show any significant association with breast cancer risk (Table 4-8, Table 4-9, Table 4-10). LH levels were associated with reduced breast cancer risk after adjustment for androstenedione and FSH did not show any significant association with breast cancer risk (Table 4-11). Neither ER-α, ER-β nor AR SB showed any significant association with breast cancer risk (Table 4-12).

Table 4-8: Association of oestrogens with risk of breast cancer - all cases.

Hormones		Oestradiol	Free oestradiol	Oestrone
		OR* (95% CI)	OR* (95% CI)	OR* (95% CI)
Not adjusted		1.080 (0.700-1.650) p=0.723	1.207 (0.791-1.827) p=0.378	1.438 (0.956-2.154) p=0.079
Hormones		**Oestradiol**	**Free oestradiol**	**Oestrone**
		OR (95% CI)**	**OR** (95% CI)**	**OR** (95% CI)**
Adjusted	Oestradiol		1.467(0.785-2.753) p=0.229	1.390 (0.919-2.093) p=0.116
	Free oestradiol	0.808 (0.420-1.525) p=0.515		1.437 (0.951-2.161) p=0.083
	Oestrone	1.066 (0.689-1.631) p=0.772	1.259 (0.824-1.909) p=0.281	
	Androstenedione	0.971 (0.619-1.505) p=0.898	1.133 (0.731-1.740) p=0.572	1.451 (0.957-2.189) p=0.077
	Testosterone	0.808 (0.506-1.272) p=0.364	0.927 (0.58-1.446) p=0.741	1.316 (0.865-1.991) p=0.196
	Free testosterone	0.915 (0.557-1.432) p=0.703	0.989 (0.616-1.568) p=0.962	1.375 (0.908-2.072) p=0.129
	DHEAS	1.075 (0.694-1.647) p=0.743	1.205 (0.786-1.831) p=0.387	1.435 (0.954-2.150) p=0.080
	SHBG	1.060 (0.684-1.625) p=0.790	1.096 (0.714-1.669) p=0.670	1.437 (0.951-2.162) p=0.082
	Progesterone	1.055 (0.680-1.619) p=0.809	1.231 (0.804-1.871) p=0.333	1.435 (0.953-2.152) p=0.082
	LH	1.056 (0.683-1.615) p=0.804	1.158 (0.756-1.759) p=0.494	1.425 (0.946-2.135) p=0.088
	FSH	1.091 (0.704-1.673) p=0.693	1.224 (0.797-1.866) p=0.351	1.446 (0.960-2.166) p=0.075
	ER-α SB	1.031 (0.665-1.580) p=0.889	1.205 (0.787-1.829) p=0.385	1.405 (0.930-2.111) p=0.102
	ER-β SB	1.047 (0.676-1.604) p=0.833	1.224 (0.800-1.857) p=0.345	1.427 (0.944-2.147) p=0.089
	AR SB	1.013 (0.651-1.557) p=0.954	1.218 (0.794-1.852) p=0.361	1.429 (0.947-2.147) p=0.087

OR values for top-bottom classification are based on controls only. *OR - for single hormones not adjusted in relation to other hormones or SB, only age adjusted. **OR - for hormones adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Table 4-9: Association of androgens with risk of breast cancer - all cases.

Hormones		Androstenedione	Testosterone	Free testosterone	DHEAS
		OR* (95% CI)	OR* (95% CI)	OR* (95% CI)	OR* (95% CI)
Not adjusted		1.854 (1.24-2.761) **p=0.002**	2.238 (1.512-3.317) **p<0.0001**	1.637 (1.090-2.453) **p=0.017**	1.048 (0.665-1.636) p=0.837
Hormones		**Androstenedione**	**Testosterone**	**Free testosterone**	**DHEAS**
		OR (95% CI)**	**OR** (95% CI)**	**OR** (95% CI)**	**OR** (95% CI)**
Adjusted	Oestradiol	1.802 (1.193-2.718) **p=0.005**	2.367 (1.565-3.591) **p<0.0001**	1.682 (1.095-2.578) p=0.703	1.044 (0.658-1.640) p=0.852
	Free oestradiol	1.751 (1.157-2.646) **p=0.008**	2.289 (1.513-3.472) **p<0.0001**	1.653 (1.054-2.590) **p=0.028**	1.015 (0.639-1.594) p=0.950
	Oestrone	1.826 (1.220-2.731) **p=0.003**	2.194 (1.478-3.260) **p≤0.0001**	1.663 (1.104-1.844) **p=0.015**	1.069 (0.678-1.671) p=0.770
	Androstenedione		1.906 (1.189-3.057) **p=0.007**	1.300 (0.826-2.033) p=0.253	0.785 (0.473-1.283) p=0.340
	Testosterone	1.259 (0.776-2.028) p=0.346		0.946 (0.550-1.600) p=0.837	0.589 (0.340-0.997) **p=0.050**
	Free testosterone	1.661 (1.071-2.571) **p=0.022**	2.338 (1.409-3.910) **p=0.001**		0.837 (0.507-1.366) p=0.483
	DHEAS	2.006 (1.304-3.090) **p=0.001**	2.819 (1.785-4.499) **p<0.0001**	1.743 (1.120-2.714) **p=0.013**	
	SHBG	1.814 (1.212-2.710) **p=0.003**	2.165 (1.457-3.120) **p<0.0001**	1.462 (0.967-2.204) *p=0.070*	1.022 (0.645-1.603) p=0.924
	Progesterone	1.923 (1.239-2.988) **p=0.004**	2.612 (1.666-4.128) **p<0.0001**	1.699 (1.099-2.624) **p=0.017**	0.941 (0.516-1.693) p=0.840
	LH	1.910 (1.277-2.855) **p=0.002**	2.298 (1.548-3.417) **p<0.0001**	1.636 (1.088-2.454) **p=0.018**	1.089 (0.688-1.706) p=0.713
	FSH	1.860 (1.246-2.166) **p=0.002**	2.243 (1.513-3.329) **p<0.0001**	1.637 (1.085-2.463) **p=0.018**	1.047 (0.662-1.637) p=0.843
	ER-α SB	1.769 (1.181-2.65) **p=0.005**	2.193 (1.477-3.260) **p<0.0001**	1.578 (1.408-2.371) **p=0.028**	1.070 (0.677-1.674) p=0.768
	ER-β SB	1.846 (1.234-2.577) **p=0.003**	2.194 (1.479-3.259) **p<0.0001**	1.594 (1.058-2.394) **p=0.025**	1.085 (0.689-1.691) p=0.721
	AR SB	1.808 (1.207-2.703) **p=0.004**	2.148 (1.446-3.191) **p<0.0001**	1.592 (1.056-2.390) **p=0.025**	1.003 (0.634-1.571) p=0.988

OR values for top-bottom classification are based on controls only. *OR - for single hormones not adjusted in relation to other hormones or SB, only age adjusted. **OR - for hormones adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Table 4-10: Association of sex hormone-binding globulin and progesterone with risk of breast cancer - all cases.

Hormones		SHBG	Progesterone
		OR* (95% CI)	OR* (95% CI)
Not adjusted		0.430 (0.245-0.720) **p=0.002**	1.124 (0.720-1.739) p=0.602
Hormones		**SHBG**	**Progesterone**
		OR (95% CI)**	**OR** (95% CI)**
Adjusted	Oestradiol	0.464 (0.267-0.772) **p=0.004**	1.077 (0.683-1.681) p=0.748
	Free oestradiol	0.464 (0.267-0.776) **p=0.005**	1.053 (0.667-1.643) p=0.823
	Oestrone	0.401 (0.226-0.679) **p=0.001**	1.198 (0.773-1.844) p=0.413
	Androstenedione	0.422 (0.236-0.716) **p=0.002**	0.841 (0.509-1.368) p=0.491
	Testosterone	0.4291 (0.240-0.730) **p=0.002**	0.655 (0.387-1.090) p=0.109
	Free testosterone	0.414 (0.229-0.713) **p=0.002**	0.890 (0.548-1.429) p=0.635
	DHEAS	0.430 (0.245-0.720) **p=0.002**	1.168 (0.653-2.081) p=0.597
	SHBG		1.115 (0.712-1.732) p=0.630
	Progesterone	0.456 (0.263-0.758) **p=0.003**	
	LH	0.432 (0.246-0.724) **p=0.002**	1.146 (0.732-1.776) p=0.547
	FSH	0.428 (0.244-0.718) **p=0.002**	1.114 (0.711-1.728 p=0.633
	ER-α SB	0.445 (0.254-0.747) **p=0.003**	1.215 (0.781-1.875) p=0.383
	ER-β SB	0.425 (0.241-0.715) **p=0.001**	1.068 (0.683-1.654) p=0.771
	AR SB	0.409 (0.231-0.692) **p=0.001**	1.137 (0.730-1.753) p=0.566

OR values for top-bottom classification are based on controls only. *OR - for single hormones not adjusted in relation to other hormones or SB, only age adjusted. **OR - for hormones adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydro- epiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Table 4-11: Association of gonadotrophins with risk of breast cancer - all cases.

Hormones		LH	FSH
		OR* (95% CI)	OR* (95% CI)
Not adjusted		0.652 (0.395-1.051) p=0.086	0.981 (0.630-1.509) p=0.932
Hormones		LH	FSH
		OR** (95% CI)	OR** (95% CI)
Adjusted	Oestradiol	0.660 (0.399-1.066) p=0.097	1.075 (0.690-1.657) p=0.745
	Free oestradiol	0.662 (0.399-1.071) p=0.100	1.088 (0.697-1.680) p=0.707
	Oestrone	0.658 (0.398-1.063) p=0.094	1.104 (0.712-1.695) p=0.654
	Androstenedione	0.556 (0.327-0.916) **p=0.025**	1.053 (0.672-1.631) p=0.820
	Testosterone	0.616 (0.327-1.003) p=0.057	1.030 (0.654-1.603) p=0.896
	Free testosterone	0.663 (0.400-1.071) p=0.100	0.990 (0.633-1.557) p=0.998
	DHEAS	0.647 (0.391-1.045) p=0.081	0.985 (0.631-1.519) p=0.948
	SHBG	0.659 (0.398-1.066) p=0.096	1.041 (0.665-1.613) p=0.858
	Progesterone	0.646 (0.391-1.042) p=0.079	0.926 (0.588-1.438) p=0.737
	LH		1.203 (0.737-1.951) p=0.456
	FSH	0.598 (0.344-1.02) p=0.062	
	ER-α SB	0.659 (0.399-1.063) p=0.095	0.947 (0.602-1.468) p=0.809
	ER-β SB	0.652 (0.395-1.050) p=0.085	0.917 (0.584-1.418) p=0.700
	AR SB	0.656 (0.396-1.058) p=0.091	0.923 (0.588-1.429) p=0.723

OR values for top-bottom classification are based on controls only. *OR - for single hormones not adjusted in relation to other hormones or SB, only age adjusted. **OR - for hormones adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Table 4-12: Association of serum bioactivity of steroid receptors with risk of breast cancer - all cases.

SB		ER-α	ER-β	AR
		OR* (95% CI)	**OR* (95% CI)**	**OR* (95% CI)**
Not adjusted		1.371 (0.905-2.603) p=0.133	1.028 (0.665-1.570) p=0.900	1.125 (0.733-1.711) p=0.584
SB		**ER-α**	**ER-β**	**AR**
		OR (95% CI)**	**OR** (95% CI)**	**OR** (95% CI)**
Adjusted	Oestradiol	1.332 (0.876-2.011) p=0.175	1.047 (0.676-1.604) p=0.833	1.138 (0.737-1.739) p=0.554
	Free oestradiol	1.326 (0.872-2.003) p=0.183	1.038 (0.670-1.592) p=0.345	1.115 (0.720-1.708) p=0.361
	Oestrone	1.305 (0.858-1.971) p=0.209	0.999 (0.643-1.534) p=0.995	1.115 (0.724-1.701) p=0.616
	Androstenedione	1.291 0.858-1.971) p=0.233	0.995 (0.637-1.534) p=0.980	1.162 (0.753-1.777) p=0.491
	Testosterone	1.295 (0.847-1.967) p=0.228	0.998 (0.641-1.537) p=0.995	1.109 (0.715-1.701) p=0.640
	Free testosterone	1.316 (0.865-1.988) p=0.195	1.016 (0.655-1.558) p=0.944	1.099 (0.711-1.681) p=0.668
	DHEAS	1.372 (0.906-2.066) p=0.131	1.029 (0.666-1.571) p=0.897	1.125 (0.733-1.711) p=0.584
	SHBG	1.309 (0.862-1.977) p=0.202	0.936 (0.602-1.437) p=0.765	1.092 (0.709-1.666) p=0.685
	Progesterone	1.343 (0.884-2.028) p=0.163	1.030 (0.666-1.575) p=0.894	1.146 (0.746-1.744) p=0.529
	LH	1.376 (0.908-2.073) p=0.129	1.028 (0.665-1.573) p=0.899	1.092 (0.710-1.663) p=0.684
	FSH	1.366 (0.902-2.058) p=0.137	1.025 (0.663-1.567) p=0.909	1.122 (0.731-1.706) p=0.594
	ER-α SB		0.880 (0.536-1.426) p=0.608	1.021 (0.642-1.607) p=0.928
	ER-β SB	1.407 (0.879-2.246) p=0.152		1.161 (0.739-1.808) p=0.514
	AR SB	1.325 (0.845-2.065) p=0.216	1.000 (0.631-1.565) p=0.999	

OR values for top-bottom classification are based on controls only. *OR - for single SB of steroid receptors not adjusted in relation to other hormones or SB, only age adjusted. **OR - for single SB of steroid receptors adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

For those women who had given a sample ≤2 years before diagnosis serum levels in the top quintile of androstenedione, testosterone, free testosterone and oestrone were significantly associated with a 1.823 (95% CI: 1.095-2.933), 2.240 (95% CI: 1.368-3.639), 1.705 (95% CI: 1.018-2.814) and 1.777 (95% CI: 1.076-2.893) fold risk for breast cancer, respectively. Testosterone remained significant after adjustment for all other hormones. However, the association of androstenedione and oestrone with breast cancer risk was not statistically significant after adjustment for testosterone (Table 4-13, Table 4-14). In addition, women who had serum levels in the top quintile of SHBG had a reduced risk of breast cancer (0.347; 95% CI: 0.150-0.705; P0.007) which remained significant after all adjustments (Table 4-15). Progesterone (Table 4-15), FSH and LH (Table 4-16), ER-α, ER-β and AR SB (Table 4-17) did not show any significant association with breast cancer risk.

Table 4-13: Association of oestrogens with risk of breast cancer - cases that gave a sample less than 2 years before diagnosis.

Hormones		Oestradiol	Free oestradiol	Oestrone
		OR* (95% CI)	OR* (95% CI)	OR* (95% CI)
Not adjusted		0.891 (0.492-1.548) p=0.692	1.134 (0.652-1.916) p=0.645	1.777 (1.076-2.893) **p=0.022**
Hormones		**Oestradiol**	**Free oestradiol**	**Oestrone**
		OR (95% CI)**	**OR** (95% CI)**	**OR** (95% CI)**
Adjusted	Oestradiol		1.551 (0.713-3.311) p=0.260	1.738 (1.045-2.847) **p=0.030**
	Free oestradiol	0.645 (0.281-1.427) p=0.288		1.754 (1.032-2.938) **p=0.035**
	Oestrone	0.869 (0.477-1.515) p=0.631	1.152 (0.659-1.954) p=0.609	
	Androstenedione	0.825 (0.448-1.458) p=0.522	1.056 (0.563-1.826) p=0.849	1.785 (1.068-2.941) **p=0.024**
	Testosterone	0.674 (0.60-1.208) p=0.199	0.862 (0.476-1.511) p=0.614	1.616 (0.966-2.661) p=0.063
	Free testosterone	0.755 (0.406-1.345) p=0.200	0.916 (0.149-1.633) p=0.722	1.727 (1.040-2.879) **p=0.032**
	DHEAS	0.878 (0.483-1.531) p=0.658	1.122 (0.641-1.905) p=0.678	1.767 (1.070-2.879) **p=0.024**
	SHBG	0.875 (0.481-1.526) p=0.648	1.017 (0.581-1.727) p=0.951	1.789 (1.077-2.932) **p=0.022**
	Progesterone	0.845 (0.463-1.479 p=0.569	1.095 (0.624-1.865) p=0.744	1.771 (1.072-2.887) **p=0.023**
	LH	0.881 (0.486-1.531) p=0.663	1.106 (0.634-1.875) p=0.714	1.766 (1.070-2.877) **p=0.024**
	FSH	0.904 (0.498-1.575) p=0.730	1.161 (0.663-1.974) p=0.591	1.791 (1.084-2.920) **p=0.021**
	ER-α SB	0.835 (0.455-1.464) p=0.543	1.092 (0.621-1.858) p=0.753	1.737 (1.046-2.843) **p=0.030**
	ER-β SB	0.843 (0.459-1.480) p=0.566	1.122 (0.637-1.914) p=0.680	1.816 (1.089-2.986) **p=0.020**
	AR SB	0.835 (0.455-1.464) p=0.543	1.090 (0.618-1.860) p=0.759	1.739 (1.048-2.844) **p=0.029**

OR values for top-bottom classification are based on controls only. *OR - for single hormones not adjusted in relation to other hormones or SB, only age adjusted. **OR - for hormones adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Table 4-14: Association of androgens with risk of breast cancer - cases that gave a sample less than 2 years before diagnosis.

Hormones		Androstenedione	Testosterone	Free testosterone	DHEAS
		OR* (95% CI)	OR* (95% CI)	OR* (95% CI)	OR* (95% CI)
Not adjusted		1.823 (1.095-2.933) **p=0.019**	2.240 (1.368-3.639) **p=0.001**	1.705 (1.018-2.814) **p=0.039**	1.103 (0.614-1.924) p=0.737
Hormones		**Androstenedione**	**Testosterone**	**Free testosterone**	**DHEAS**
		OR (95% CI)**	**OR** (95% CI)**	**OR** (95% CI)**	**OR** (95% CI)**
Adjusted	Oestradiol	1.827 (1.081-3.046) **p=0.022**	2.457 (1.469-4.069) **p<0.0001**	1.842 (1.077-3.111) **p=0.024**	1.137 (0.630-1.999) p=0.660
	Free oestradiol	1.754 (1.032-2.938) **p=0.035**	2.338 (1.390-3.912) **p=0.001**	1.786 (1.017-3.099) **p=0.041**	1.089 (0.601-1.917) p=0.773
	Oestrone	1.811 (1.083-2.989) **p=0.021**	2.150 (1.306-3.511) **p=0.002**	1.733 (1.031-2.872) **p=0.035**	1.145 (0.639-1.999) p=0.641
	Androstenedione		1.831 (1.009-3.287) **p=0.044**	1.328 (0.71-2.333) p=0.331	0.858 (0.448-1.586) p=0.632
	Testosterone	1.284 (0.698-2.319) p=0.413		1.018 (0.517-1.960) p=0.959	0.650 (0.329-1.246) p=0.204
	Free testosterone	1.655 (0.951-2.873) p=0.070	2.259 (1.194-4.256) **p=0.001**		0.886 (0.467-1.632) p=0.704
	DHEAS	1.912 (1.109-3.260) **p=0.018**	2.690 (1.528-4.729) **p<0.0001**	1.779 (1.017-3.072) **p=0.041**	
	SHBG	1.771 (1.059-2.922) **p=0.027**	2.122 (1.290-3.463) **p=0.003**	1.473 (0.874-2.448) p=0.139	1.047 (0.580-1.837) p=0.875
	Progesterone	1.843 (1.055-3.185) **p=0.030**	2.463 (1.412-4.282) **p=0.001**	1.682 (0.967-2.886) *p=0.061*	0.901 (0.419-1.886) p=0.786
	LH	1.847 (1.108-3.038) **p=0.017**	2.263 (1.381-3.682) **p=0.001**	1.705 (1.017-2.816) **p=0.039**	1.128 (0.627-1.975) p=0.679
	FSH	1.838 (1.102-3.024) **p=0.018**	2.264 (1.379-3.692) **p=0.001**	1.713 (1.017-2.844) **p=0.039**	1.111 (1.017-2.844) p=0.717
	ER-α SB	1.781 (1.063-2.939) **p=0.026**	2.199 (1.336-3.591) **p=0.002**	1.629 (0.967-2.777) *p=0.062*	1.124 (0.626-1.964) p=0.687
	ER-β SB	1.825 (1.087-3.019) **p=0.021**	2.194 (1.332-3.585) **p=0.002**	1.672 (0.991-2.777) **p=0.050**	1.174 (0.658-2.041) p=0.577
	AR SB	1.782 (1.064-2.940) **p=0.025**	2.156 (1.312-3.514) **p=0.002**	1.630 (0.967-2.702) *p=0.061*	1.102 (0.614-1.920) p=0.738

OR values for top-bottom classification are based on controls only. *OR - for single hormones not adjusted in relation to other hormones or SB, only age adjusted. **OR - for hormones adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Table 4-15: Association of sex hormone-binding globulin and progesterone with risk of breast cancer - cases that gave a sample less than 2 years before diagnosis.

Hormones		SHBG	Progesterone
		OR* (95% CI)	OR* (95% CI)
Not adjusted		0.347 (0.150-0.705) **p=0.007**	1.258 (0.715-2.159) p=0.413
Hormones		**SHBG**	**Progesterone**
		OR (95% CI)**	**OR** (95% CI)**
Adjusted	Oestradiol	0.352 (0.152-0.715) **p=0.007**	1.296 (0.731-2.242) p=0.363
	Free oestradiol	0.349 (0.150-0.714) **p=0.007**	1.245 (0.702-2.156) p=0.443
	Oestrone	0.296 (0.121-0.625) **p=0.003**	1.212 (0.688-2.084) p=0.495
	Androstenedione	0.321 (0.130-0.677) **p=0.006**	0.967 (0.511-1.777) p=0.915
	Testosterone	0.324 (0.131-0.685) **p=0.006**	0.801 (0.420-1.485) p=0.490
	Free testosterone	0.279 (0.105-0.617) **p=0.004**	1.039 (0.565-1.864) p=0.899
	DHEAS	0.348 (0.150-0.708) **p=0.006**	1.343 (0.643-2.750) p=0.425
	SHBG		1.244 (0.704-2.148) p=0.441
	Progesterone	0.348 (0.150-0.708) **p=0.007**	
	LH	0.348 (0.150-0.708) **p=0.007**	1.278 (0.725-2.2197) p=0.384
	FSH	0.343 (0.148-0.699) **p=0.006**	1.253 (0.710-2.158) p=0.425
	ER-α SB	0.351 (0.151-0.716) **p=0.008**	1.183 (0.667-2.044) p=0.554
	ER-β SB	0.316 (0.136-0.647) **p=0.003**	1.164 (0.657-2.006) p=0.593
	AR SB	0.347 (0.149-0.705) **p=0.007**	1.162 (0.657-2.000) p=0.595

OR values for top-bottom classification are based on controls only. *OR - for single hormones not adjusted in relation to other hormones or SB, only age adjusted. **OR - for hormones adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Table 4-16: Association of gonadotrophins with risk of breast cancer - cases that gave a sample less than 2 years before diagnosis.

Hormones		LH	FSH
		OR* (95% CI)	OR* (95% CI)
Not adjusted		0.768 (0.408-1.376) p=0.391	1.067 (0.608-1.813) p=0.814
Hormones		**LH**	**FSH**
		OR (95% CI)**	**OR** (95% CI)**
Adjusted	Oestradiol	0.775 (0.411-1.391) p=0.409	1.143 (0.654-1.934) p=0.629
	Free oestradiol	0.780 (0.413-1.404) p=0.423	1.165 (0.665-1.985) p=0.582
	Oestrone	0.780 (0.413-1.402) p=0.422	1.230 (0.709-2.078 p=0.448
	Androstenedione	0.617 (0.308-1.155) p=0.148	1.113 (0.624-1.920) p=0.709
	Testosterone	0.749 (0.395-1.395) p=0.355	1.118 (0.626-1.934) p=0.697
	Free testosterone	0.785 (0.417-1.410) p=0.435	1.036 (0.575-1.801) p=0.903
	DHEAS	0.760 (0.403-1.363) p=0.374	1.078 (0.613-1.837) p=0.787
	SHBG	0.777 (0.412-1.398) p=0.417	1.140 (0.645-1.952) p=0.642
	Progesterone	0.754 (0.400-1.352) p=0.360	0.964 (0.537-1.668) p=0.899
	LH		1.231 (0.664-2.221) p=0.498
	FSH	0.698 (0.349-1.333) p=0.290	
	ER-α SB	0.781 (0.415-1.400) p=0.423	0.959 (0.535-1.653) p=0.883
	ER-β SB	0.773 (0.410-1.387) p=0.404	0.931 (0.519-1.606) p=0.804
	AR SB	0.768 (0.408-1.378) p=0.392	0.945 (0.528-1.628) p=0.844

OR values for top-bottom classification are based on controls only. *OR - for single hormones not adjusted in relation to other hormones or SB, only age adjusted. **OR - for hormones adjusted for the other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Table 4-17: Association of serum bioactivity of sex steroid receptors with risk of breast cancer - cases that gave a sample less than 2 years before diagnosis.

SB		ER-α	ER-β	AR
		OR* (95% CI)	OR* (95% CI)	OR* (95% CI)
Not adjusted		1.021 (0.576-1.746) p=0.941	0.651 (0.339-1.176) p=0.173	0.995 (0.561-1.704) p=0.986
SB		**ER-α**	**ER-β**	**AR**
		OR (95% CI)**	**OR** (95% CI)**	**OR** (95% CI)**
Adjusted	Oestradiol	0.965 (0.538-1.665) p=0.900	0.665 (0.345-1.205) p=0.198	1.044 (0.588-1.792) p=0.878
	Free oestradiol	0.959 (0.538-1.665) p=0.884	0.660 (0.342-1.196) p=0.198	1.044 (0.588-1.792) p=0.921
	Oestrone	0.928 (0.516-1.606) p=0.796	0.612 (0.316-1.116) p=0.125	0.989 (0.556-1.700) p=0.970
	Androstenedione	0.906 (0.497-1.584) p=0.736	0.607 (0.307-1.121) p=0.128	1.070 (0.600-1.845) p=0.812
	Testosterone	0.914 (0.506-1.589) p=0.757	0.636 (0.329-1.159) p=0.156	1.019 (0.571-1.757) p=0.946
	Free testosterone	0.947 (0.527-1.636) p=0.849	0.638 (0.339-1.180) p=0.157	1.007 (0.566-1.730) p=0.980
	DHEAS	1.022 (0.577-1.748) p=0.938	0.653 (.0339-0.180) p=0.177	0.994 (0.560-1.701) p=0.983
	SHBG	0.938 (0.527-1.612) p=0.821	0.556 (0.288-1.012) p=0.065	0.932 (0.523-1.603) p=0.805
	Progesterone	0.963 (0.537-1.661) p=0.895	0.652 (0.339-1.179) p=0.176	1.017 (0.573-1.743) p=0.951
	LH	1.030 (0.581-1.763) p=0.917	0.652 (0.339-1.179) p=0.176	0.978 (0.551-1.677) p=0.938
	FSH	1.019 (0.574-1.744) p=0.948	0.649 (0.337-1.173) p=0.171	0.995 (0.561-1.703) p=0.985
	ER-α SB		0.611 (0.298-1.83) p=0.159	1.024 (0.558-1.820) p=0.936
	ER-β SB	1.170 (0.616-2.153) p=0.621		1.142 (0.632-1.998) p=0.649
	AR SB	0.946 (0.510-1.691) p=0.855	0.637 (0.325-1.174) p=0.165	

OR values for top-bottom classification are based on controls only. *OR - for single SB of sex steroid receptors not adjusted in relation to other hormones or SB, only age adjusted. **OR - for SB sex steroid receptors adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

For those women who had given a sample >2 years before diagnosis serum levels in the top quintile of oestradiol, free oestradiol and oestrone did not show any significant association with breast cancer risk. Serum levels in the top quintile androstenedione and testosterone were significantly associated with 1.868 (95% CI: 1.120-3.073) and 2.218 (95% CI: 1.341-3.634) fold risk for breast cancer, respectively. Whereas testosterone remained significant after adjustment for all other hormones, androstenedione did not retain significance after adjustment for testosterone (Table 4-19). In addition women with serum ER-α bioactivity (≥104.359 pg/ml) in the top quintile had a 1.791 (95% CI: 1.070-2.951; $P<0.05$) fold breast cancer risk. This association remained statistically significant after adjustment for other hormones and AR SB. No association was shown between breast cancer risk and ER-β and AR SB (Table 4-22). Other hormones tested did not show any significant association with breast cancer risk.

Table 4-18: Association of oestrogens with risk of breast cancer - cases that gave a sample more than 2 years before diagnosis.

Hormones		Oestradiol	Free oestradiol	Oestrone
		OR* (95% CI)	OR* (95% CI)	OR* (95% CI)
Not adjusted		1.294 (0.747-2.184) p=0.343	1.284 (0.742-2.165) p=0.358	1.116 (0.634-1.901) p=0.695
Hormones		**Oestradiol**	**Free oestradiol**	**Oestrone**
		OR (95% CI)**	**OR** (95% CI)**	**OR** (95% CI)**
Adjusted	Oestradiol		1.331 (0.594-2.923) p=0.480	1.076 p=0.796
	Free oestradiol	1.045 (0.461-2.313) p=0.914		1.161 (0.658-1.985) p=0.695
	Oestrone	1.290 (0.744-2.178) p=0.351	1.380 (0.801-2.322) p=0.233	
	Androstenedione	1.148 (0.652-1.966) p=0.622	1.206 (0.687-2.061) p=0.503	1.129 (0.639-1.932) p=0.666
	Testosterone	0.993 (0.552-1.734) p=0.981	1.008 (0.561-1.758) p=0.978	1.038 (0.580-1.793) p=0.897
	Free testosterone	1.137 (0.637-1.975) p=0.654	1.089 (0.591-1.950) p=0.780	1.049 (0.589-1.805) p=0.866
	DHEAS	1.304 (0.748-2.213) p=0.336	1.296 (0.744-2.201) p=0.346	1.117 (0.634-1.904) p=0.693
	SHBG	1.276 (0.735-2.158) p=0.373	1.192 (0.684-2.205) p=0.524	1.118 (0.634-1.910) p=0.691
	Progesterone	1.300 (0.746-2.208) p=0.342	1.379 (0.798-2.329) p=0.238	1.115 (0.633-1.903) p=0.696
	LH	1.249 (0.719-2.113) p=0.416	1.198 (0.689-2.031) p=0.510	1.099 (0.624-1.876) p=0.736
	FSH	1.295 (0.743-2.201) p=0.348	1.285 (0.735-2.193) p=0.366	1.113 (0.63201.899) p=0.701
	ER-α SB	1.252 (0.720-2.120) p=0.413	1.316 (0.761-2.218) p=0.313	1.085 (0.614-1.858) p=0.771
	ER-β SB	1.268 (0.731-2.144) p=0.384	1.325 (0.767-2.232) p=0.300	1.073 (0.606-1.840) p=0.803
	AR SB	1.216 (0.695-2.067) p=0.480	1.363 (0.790-2.298) p=0.254	1.139 (0.646-1.945) p=0.641

OR values for top-bottom classification are based on controls only. *OR - for single hormones not adjusted in relation to other hormones or SB, only age adjusted. **OR - for hormones adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Table 4-19: Association of androgens with risk of breast cancer - cases that gave a sample more than 2 years before diagnosis.

Hormones		Androstenedione	Testosterone	Free testosterone	DHEAS
		OR* (95% CI)	OR* (95% CI)	OR* (95% CI)	OR* (95% CI)
Not adjusted		1.868 (1.120-3.073) **p=0.015**	2.218 (1.341-3.634) **p=0.001**	1.558 (0.916-2.599) p=0.095	0.990 (0.535-1.762) p=0.972
Hormones		**Androstenedione**	**Testosterone**	**Free testosterone**	**DHEAS**
		OR (95% CI)**	**OR** (95% CI)**	**OR** (95% CI)**	**OR** (95% CI)**
Adjusted	Oestradiol	1.756 (1.036-2.932) **p=0.033**	2.215 (1.304-3.733) **p=0.003**	1.490 (0.82-2.555) p=0.153	0.947 (0.508-1.700) p=0.859
	Free oestradiol	1.735 (1.020-2.909) **p=0.039**	2.213 (1.333-3.622) **p=0.002**	1.499 (0.832-2.652) p=0.169	0.935 (0.500-1.680) 0.826
	Oestrone	1.801 (1.074-2.977) **p=0.023**	2.207 (1.333-3.622) **p=0.001**	1.574 (0.923-2.632) p=0.089	0.983 (0.531-1.752) p=0.954
	Androstenedione		1.959 (1.086-3.505) **p=0.024**	1.282 (0.719-2.236) p=0.389	0.748 (0.385-1.393) p=0.372
	Testosterone	1.262 (0.684-2.281) p=0.448		0.906 (0.458-1.745) p=0.770	0.539 (0.263-0.972) p=0.082
	Free testosterone	1.645 (0.941-2.832) p=0.076	2.350 (1.245-4.414) **p=0.008**		0.781 (0.397-1.479) p=0.461
	DHEAS	2.038 (1.182-3.478) **p=0.010**	2.895 (1.618-5.173) **p≤0.0001**	1.705 (0.954-3.001) p-=0.067	
	SHBG	1.849 (1.106-3.050) **p=0.017**	2.161 (1.305-3.549) **p=0.002**	1.439 (0.840-2.418) p=0.176	0.969 (0.522-1.732) p=0.917
	Progesterone	1.963 (1.123-3.395) **p=0.016**	2.675 (1.514-4.711) **p<0.0001**	1.714 (0.971-2.977) p=0.058	0.984 (0.457-2.047) p=0.966
	LH	1.891 (1.132-3.119) **p=0.013**	2.220 (1.340-3.645) **p=0.002**	1.532 (0.898-2.561) p=0.109	1.030 (0.555-1.841) p=0.923
	FSH	1.867 (1.118-3.075) **p=0.015**	2.218 (1.336-3.652) **p=0.002**	1.551 (0.905-2.606) p=0.103	0.977 (0.527-1.745) p=0.938
	ER-α SB	1.751 (1.042-2.900) **p=0.031**	2.178 (1.312-3.582) **p=0.002**	1.527 (0.894-2.555) p=0.113	1.000 (0.538-1.789) p=1.000
	ER-β SB	1.858 (1.111-3.065) **p=0.016**	2.165 (1.306-3.557) **p=0.002**	1.520 (0.891-2.540) p=0.166	0.983 (0.530-1.753) p=0.954
	AR SB	1.819 (1.084-3.008) **p=0.021**	2.125 (1.279-3.495) **p=0.003**	1.543 (0.905-2.578) p=0.103	0.898 (0.479-1.614) p=0.726

OR values for top-bottom classification are based on controls only. *OR - for single hormones not adjusted in relation to other hormones or SB, only age adjusted. **OR - for hormones adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Table 4-20: Association of sex hormone-binding globulin and progesterone with risk of breast cancer - cases that gave a sample more than 2 years before diagnosis.

Hormones		SHBG	Progesterone
		OR* (95% CI)	**OR* (95% CI)**
Not adjusted		0.518 (0.250-0.982) p=0.057	0.992 (0.539-1.753) p=0.978
Hormones		SHBG	**Progesterone**
		OR** (95% CI)	**OR** (95% CI)**
Adjusted	Oestradiol	0.589 (0.292-1.101) p=0.115	0.876 (0.466-1.577) p=0.668
	Free oestradiol	0.596 (0.294-1.121) p=0.126	0.870 (0.463-1.566) p=0.652
	Oestrone	0.513 (0.248-0.973) p=0.054	1.183 (0.664-2.047) p=0.556
	Androstenedione	0.522 (0.251-0.994) p=0.061	0.741 (0.380-1.384) p=0.361
	Testosterone	0.551 (0.265-1.052) p=0.087	0.550 (0.271-1.066) p=0.085
	Free testosterone	0.563 (0.270-1.080) p=0.101	0.741 (0.379-1.385) p=0.362
	DHEAS	0.517 (0.250-0.981) p=0.056	1.001 (0.469-2.067) p=0.997
	SHBG		0.987 (0.536-1.650) p=0.966
	Progesterone	0.573 (0.284-1.068) p=0.096	
	LH	0.525 (0.253-0.997) p=0.063	1.003 (0.544-1.779) p=0.993
	FSH	0.521 (0.252-0.990) p=0.060	0.977 (0.529-1.735) p=0.938
	ER-α SB	0.547 (0.264-1.043) p=0.082	1.253 (0.700-2.181) p=0.436
	ER-β SB	0.543 (0.261-1.037) p=0.080	0.972 (0.528-1.719) p=0.924
	AR SB	0.474 (0.222-0.918) **p=0.037**	1.108 (0.616-1.930) p=0.723

OR values for top-bottom classification are based on controls only. *OR - for single hormones not adjusted in relation to other hormones or SB, only age adjusted. **OR - for hormones adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Table 4-21: Association of gonadotrophins with risk of breast cancer - cases that gave a sample more than 2 years before diagnosis.

Hormones		LH	FSH
		OR* (95% CI)	OR* (95% CI)
Not adjusted		0.535 (0.256-1.029) p=0.075	0.889 (0.483-1.565) p=0.694
Hormones		**LH**	**FSH**
		OR (95% CI)**	**OR** (95% CI)**
Adjusted	Oestradiol	0.545 (0.260-1.051) p=0.085	1.005 (0.549-1.772) p=0.985
	Free oestradiol	0.544 (0.259-1.053) p=0.086	1.005 (0.546-1.781) p=0.987
	Oestrone	0.535 (0.256-1.032) p=0.077	0.964 (0.528-1.691) p=0.902
	Androstenedione	0.523 (0.249-1.011) p=0.067	0.992 (0.543-1.742) p=0.979
	Testosterone	0.534 (0.254-1.035) p=0.078	1.001 (0.539-1.784) p=0.996
	Free testosterone	0.548 (0.262-1.055) p=0.087	0.967 (0.521-1.720) p=0.912
	DHEAS	0.534 (0.255-1.028) p=0.074	0.887 (0.480-1.566) p=0.689
	SHBG	0.543 (0.259-1.046) p=0.083	0.932 (0.504-1.648) p=0.814
	Progesterone	0.533 (0.255-1.025) p=0.074	0.884 (0.478-1.564) p=0.683
	LH		1.155 (0.594-2.165) p=0.660
	FSH	0.500 (0.226-1.029) p=0.072	
	ER-α SB	0.537 (0.256-1.037) p=0.079	0.931 (0.504-1.647) p=0.813
	ER-β SB	0.531 (0.254-1.022) p=0.072	0.905 (0.491-1.596) p=0.738
	AR SB	0.541 (0.258-1.044) p=0.082	0.904 (0.490-1.595) p=0.737

OR values for top-bottom classification are based on controls only. *OR - for single hormones not adjusted in relation to other hormones or SB, only age adjusted. **OR - for hormones adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Table 4-22: Association of serum bioactivity of sex steroid receptors with risk of breast cancer - cases that gave a sample more than 2 years before diagnosis.

SB		ER-α	ER-β	AR
		OR* (95% CI)	**OR* (95% CI)**	**OR* (95% CI)**
Not adjusted		1.791 (1.070-2.951) **p=0.023**	1.495 (0.882-2.482) p=0.126	1.271 (0.734-2.143) p=0.378
SB		**ER-α**	**ER-β**	**AR**
		OR (95% CI)**	**OR** (95% CI)**	**OR** (95% CI)**
Adjusted	Oestradiol	1.775 (1.059-2.931) **p=0.026**	1.513 (0.890-2.521) p=0.117	1.237 (0.707-2.105) p=0.442
	Free oestradiol	1.766 (1.052-2.918) **p=0.028**	1.494 (0.877-2.494) p=0.130	1.218 (0.694-2.076) p=0.479
	Oestrone	1.773 (1.058-2.928) **p=0.027**	1.496 (0.878-2.500) p=0.130	1.280 (0.738-2.161) p=0.366
	Androstenedione	1.746 (1.039-2.890) **p=0.032**	1.464 (0.858-2.455) p=0.152	1.281 (0.737-2.171) p=0.366
	Testosterone	1.745 (1.035-2.898) **p=0.033**	1.450 (0.848-2.428) p=0.164	1.235 (0.703-2.108) p=0.449
	Free testosterone	1.776 (1.059-2.933) **p=0.027**	1.493 (0.878-2.489) p=0.130	1.209 (0.737-2.156) p=0.494
	DHEAS	1.791 (1.070-2.951) **p=0.023**	1.496 (0.878-2.484) p=0.126	1.278 (0.737-2.156) p=0.369
	SHBG	1.726 (1.029-2.851) **p=0.035**	1.392 (0.816-2.325) p=0.214	1.240 (0.714-2.097) p=0.431
	Progesterone	1.810 (1.079-2.989) **p=0.022**	1.496 (0.882-2.487) p=0.127	1.296 (0.748-2.187) p=0.342
	LH	1.795 (1.071-2.965) **p=0.024**	1.500 (0.883-2.496) p=0.125	1.218 (0.702-2.060) p=0.471
	FSH	1.784 (1.065-2.943) **p=0.025**	1.480 (0.877-2.474) p=0.132	1.263 (0.729-2.133) p=0.392
	ER-α SB		1.197 (0.657-2.133) p=0.548	1.017 (0.558-1.801) p=0.954
	ER-β SB	1.646 (0.920-2.901) p=0.080		1.161 (0.650-2.019) p=0.604
	AR SB	1.795 (1.206-3.096) **p=0.037**	1.468 (0.841-2.513) p=0.167	

OR values for top-bottom classification are based on controls only. *OR - for SB of sex steroid receptors not adjusted in relation to other hormones or SB, only age adjusted. **OR - for SB of sex steroid hormones adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

4.3.5 Correlation among hormones and serum bioactivity

Using all samples, correlations among sex steroid hormones, SHBG, gonadotrophins and SB of the sex steroid receptors were investigated. A positive statistically significant correlation was demonstrated for free oestradiol and free testosterone with SB of ER-α, ER-β and AR. A negative statistically significant correlation was shown among SHBG with ER-α, ER-β and AR SB. All three different sex steroid hormone receptors investigated were shown to be correlated amongst them (Table 4-23). Serum concentration of free oestradiol was positively and significantly correlated with all other hormones, with correlation ranging from 0.100 to 0.897. Oestrone serum level was correlated with all other oestrogens and with testosterone. Serum concentrations of androstenedione (r ranged from 0.093 for oestrone to 0.708 for testosterone) and free testosterone (r ranged from 0.105 for oestrone to 0.875 for testosterone) were correlated with all oestrogens and androgens. Testosterone and DHEAS serum levels were correlated with serum levels of all other androgens and with oestradiol and free oestradiol. SHBG was significantly negative correlated with free oestradiol ($r=-0.524$, $p\leq0.0001$) and free testosterone ($r=-0.453$, $p\leq0.0001$) (Table 4-24). LH and FSH were demonstrated to be negatively correlated with oestrogens and SB of the sex steroid receptors. LH was shown to be negatively and FSH to be positively correlated with androgens and progesterone. A positive correlation was observed amongst FSH and LH with SHBG and between them (Table 4-25).

Table 4-23: Spearman correlation coefficients between sex steroid hormones and sex hormone-binding globulin and serum bioactivity of oestrogen and androgen receptors.

Correlation coefficients			
SB / Hormones	**ER-α SB (pg/ml)**	**ER-β SB (pg/ml)**	**AR SB (ng/ml)**
Oestradiol (pg/mL)	0.059 p=0.181	0.062 p=0.160	0.055 p=0.214
Free oestradiol (pmol/l)	0.124 **p=0.005**	0.148 **p=0.001**	0.109 **p=0.013**
Oestrone (pg/ml)	0.025 p=0.565	0.066 p=0.132	0.080 p=0.067
Androstenedione (nmol/l)	0.058 p=0.186	0.081 p=0.064	0.002 p=0.963
Testosterone (nmol/l)	0.024 p=0.592	0.051 p=0.244	0.034 p=0.443
Free Testosterone (ng/dl)	0.102 **p=0.021**	0.139 **p=0.002**	0.090 **p=0.041**
DHEAS (ug/dl)	0.020 p=0.647	0.010 p=0.814	0.012 p=0.785
SHBG (nmol/l)	-0.220 **p=0.005**	-0.242 **p<0.0001**	-0.128 **p=0.004**
Progesterone (ng/ml)	0.015 p=0.727	0.014 p=0.751	-0.007 p=0.873
ER-α SB (pg/ml)		0.507 **p<0.0001**	0.307 **p<0.0001**
ER-β SB (pg/ml)			0.330 **p<0.0001**

AR=androgen receptor; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Table 4-24: Spearman correlation coefficients among sex steroid hormones along with sex steroid hormone binding globulin.

	Correlation coefficients							
Hormones	**Oestradiol (pg/mL)**	**Free oestradiol (pmol/l)**	**Oestrone (pg/ml)**	**Androstenedione (nmol/L)**	**Testosterone (nmol/L)**	**Free Testosterone (ng/dl)**	**DHEAS (ug/dl)**	**SHBG (nmol/l)**
Free oestradiol (pmol/l)	0.897 **p<0.0001**							
Oestrone (pg/ml)	0.097 **p=0.029**	0.100 **p=0.024**						
Androstenedione (nmol/l)	0.239 **p<0.0001**	0.252 **p<0.0001**	0.093 **p=0.034**					
Testosterone (nmol/l)	0.395 **p<0.0001**	0.356 **p<0.0001**	0.075 p=0.091	0.708 **p<0.0001**				
Free Testosterone (ng/dl)	0.425 **p<0.0001**	0.554 **p<0.0001**	0.105 **p=0.018**	0.651 **p<0.0001**	0.875 **p<0.0001**			
DHEAS (ug/dl)	0.155 **p<0.0001**	0.176 **p<0.0001**	0.052 p=0.238	0.593 **p<0.0001**	0.663 **p<0.0001**	0.601 **p<0.0001**		
SHBG (nmol/l)	-0.177 p=0.062	-0.524 **p<0.0001**	-0.042 p=0.346	-0.083 p=0.062	-0.048 p=0.279	-0.453 **p<0.0001**	-0.078 p=0.075	
Progesterone (ng/ml)	0.266 **p<0.0001**	0.253 **p<0.0001**	0.045 p=0.307	0.629 **p<0.0001**	0.704 **p<0.0001**	0.625 **p<0.0001**	0.855 **p<0.0001**	-0.023 p=0.595

AR=androgen receptor; ER=oestrogen receptor; DHEAS=dehydroepiandrosterone sulphate; SHBG=sex hormone-binding globulin

Table 4-25: Spearman correlation coefficients between gonadotrophins and oestrogens or androgens or sex hormone-binding globulin along with serum bioactivity of oestrogen and androgen receptors.

Correlation coefficients		
Hormones and SB	**LH (mIU/ml)**	**FSH (mIU/ml)**
Oestradiol (pg/mL)	-0.099 **p=0.024**	-0.282 p=0.200
Free oestradiol (pmol/l)	-0.149 **p=0.001**	-0.362 **p<0.0001**
Oestrone (pg/ml)	-0.036 p=0.421	-0.051 p=0.250
Androstenedione (nmol/l)	0.052 p=0.236	-0.072 p=0.104
Testosterone (nmol/l)	0.074 p=0.091	-0.085 p=0.052
Free Testosterone (ng/dl)	-0.017 p=0.698	-0.204 **p<0.0001**
DHEAS (ug/dl)	0.113 **p=0.010**	-0.018 p=0.674
SHBG (nmol/l)	0.098 **p=0.024**	0.269 **p<0.0001**
Progesterone (ng/ml)	0.068 p=0.123	-0.049 p=0.263
FSH (mIU/ml)	0.659 **p<0.0001**	
ER-α SB (pg/ml)	-0.024 p=0.588	-0.103 p=0.019
ER-β SB (pg/ml)	-0.028 p=0.521	-0.121 **p=0.006**
AR SB (ng/ml)	-0.074 p=0.090	-0.106 p=0.016

AR=androgen receptor; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; SB=serum bioactivity; SHBG=sex hormone-binding globulin

4.3.6 Combination of hormones and serum bioactivity and their joint association with breast cancer risk

Further investigation was undertaken to examine whether joint association of hormones, hormones and steroid receptor SB and joint association of steroid receptors has a better predictive power. Additionally, analysis was carried out in relation to time of diagnosis (less and more than 2 years before diagnosis). Within the following section different pairs of hormones/SB that were significantly associated with breast cancer risk are presented and those that were not found to have significant results are included in the appendices.

Initially, joint associations between the different hormones were examined analysing all cases. When high levels (top quintiles) of the joint association of oestrogens and androgens were investigated it was shown that women with oestrone and testosterone in the top quintiles had 2.507 (95% CI: 1.495-5.738; $p \leq 0.05$) fold breast cancer risk which remained statistically significant after adjustment for other hormones and SB. Women with high levels of oestradiol and androstenedione or testosterone had 1.899 (95% CI: 1.032-3.561, $p \leq 0.05$) or 1.958 (95% CI: 1.182-3.604, $p \leq 0.05$) fold risk which did not remain statistically significant after adjustment for testosterone and androstenedione respectively (Table 4-26). After applying Bonferroni correction the significance was lost. Joint associations of androgens demonstrated that women having androstenedione and testosterone (this significance remained after bonferonni correction) or androstenedione and DHEAS levels (this significance was lost after bonferonni correction) in the top quintile had 2.567 (95% CI: 1.703-4.678; $p<0.0001$), 1.972 (95% CI: 1.106-3.483; $p \leq 0.05$) fold breast cancer risk with the latter association not remaining statistically significant after adjustment for testosterone (Table 4-

27). Women with high levels of androstenedione or testosterone with progesterone were at 1.731 (95% CI: 0.989-2.866; p≤0.05) or 1.824 (95% CI: 1.092-2.972; p≤0.05) fold risk of breast cancer which did not remain statistically significant after adjustment for testosterone and androstenedione respectively and the significance of these joint associations were lost after bonferonni correction (Table 4-28). Analysis on the joint association of androgens with gonadotrophins showed women with high levels of androstenedione and FSH (significance that remained after bonferonni correction) or testosterone and LH (significance that was lost after bonferonni correction) having 1.731 (95% CI: 0.935-4.174, p≤0.05) or 3.029 (95% CI: 1.284-5.555; p≤0.05) fold breast cancer risk. FSH and testosterone was the pair with the highest statistically significant OR that remained after adjustment for other hormones/SB. Women having these two hormones in top quintile had 5.924 (95% CI: 2.337-16.152; p<0.0001) fold breast cancer risk (Table 4-29). All other combinations were not shown to be statistically significant associated with risk of breast cancer (Appendix VI).

When women who gave samples ≤2 before diagnosis were analysed, joint association of oestrogens was shown to be statistically significant associated with breast cancer risk (OR: 2.578; 95% CI: 1.044-6.120; p≤0.05) which was lost after adjustment for testosterone and for sex steroid receptors SB (Table 4-30). When high levels (top quintiles) of the joint association of oestrogens and androgens were investigated it was shown for those women who had given samples ≤2 before diagnosis with oestrone and testosterone levels in the top quintiles to have had 3.390 (95% CI: 1.818-8.384; p≤0.05) fold risk of breast cancer (Table 4-31). High levels of androstenedione and oestrone were also shown also to be statistically significant associated with breast cancer risk but after adjustment for

ER-α and AR this significance was lost (Table 4-31). When all the above joint associations were corrected with bonferonni correction, they lost significance. Joint associations of androgens demonstrated that the only pair of hormones that remained statistically significant associated with breast cancer risk after adjust for all hormones and SB was androstenedione and testosterone (OR: 2.555; 95%CI: 1.426-4.580; p≤0.05) (Table 4-32), which also remained significantly associated with breast cancer risk after bonferonni correction. Women with high levels of androstenedione or testosterone and progesterone were at 1.928 (95%CI: 0.978-3.590; p≤0.05) or 2.025 (95%CI: 1.094-3.705; p≤0.05) fold risk of breast cancer which did not remain statistically significant after adjustment for oestrone, testosterone, SB of all steroid receptors and androstenedione respectively (Table 4-33). Finally, analysis on the joint association of androgens with gonadotrophins showed that women with high levels of testosterone and LH had 3.816 (95% CI: 1.527-7.891; p≤0.05) fold breast cancer risk (significance that was lost after bonferonni correction). The best breast cancer risk predictive pair was FSH and testosterone with women having these two hormones in the top quintile having 6.404 (95% CI: 2.620-21.648; p<0.0001) fold increased breast cancer risk (significance that remained after bonferonni correction) (Table 4-34). All other combinations were not shown to be statistically significant associated with breast cancer risk (Appendix VI).

When analysis was undertaken for those women who gave a sample >2 before diagnosis it was demonstrated that high levels (top quintiles) of the joint association of oestradiol and androstenedione were associated with 2.101 (95% CI: 1.114-5.069; p≤0.05) fold increased breast cancer risk which was not significant after adjustment for testosterone and AR SB. A significant association

with breast cancer risk was also shown for high levels of testosterone with oestradiol that remained significant after adjustment for other hormones and SB (Table 4-35). All the above joint associations though lost significance after bonferonni correction. Analysis of the joint associations of androgens demonstrated androstenedione and testosterone (OR: 2.555; 95%CI: 1.426-4.580; p≤0.05) to be the only pair of androgens that remained statistically significant associated with breast cancer risk after adjust for all hormones and SB, and after bonferonni correction (Table 4-36). Finally, analysis on the joint association of androgens with gonadotrophins demonstrated that women with testosterone and FSH levels in the top quintile to have 5.330 (95% CI: 1.456-15.174; p≤0.05) fold breast cancer risk (Table 4-37). All other combinations were not shown to be statistically significant associated with breast cancer risk (Appendix VI).

Table 4-26: Joint association of oestrogens and androgens (top quintiles) with risk of breast cancer – all cases.

Hormones		Joint association of oestrogens and androgens		
		Oestradiol		Oestrone
		Androstenedione	Testosterone	Testosterone
		OR* (95% CI)	**OR* (95% CI)**	**OR* (95% CI)**
Not adjusted		1.899 (1.010-3.567) *p=0.039*	1.958 (1.182-3.604) *p=0.016*	2.507 (1.495-5.738) *p=0.006*
		OR (95% CI)**	**OR** (95% CI)**	**OR** (95% CI)**
Adjusted	Oestradiol			2.613 (1.332-5.200) *p=0.005*
	Oestrone	1.865 (1.009-3.435) *p=0.045*	1.903 (1.095-3.303) *p=0.022*	
	Androstenedione		1.583 (0.876-2.850) p=0.125	2.131 (1.062-4.311) *p=0.033*
	Testosterone	1.196 (0.607-2.341) p=0.601		
	DHEAS	1.930 (1.029-3.611) *p=0.039*	2.014 (1.140-3.559) *p=0.015*	2.576 (1.309-5.133) *p=0.006*
	SHBG	1.863 (1.004-3.444) *p=0.047*	1.961 (1.124-3.415) *p=0.017*	2.257 (1.161-4.434) *p=0.017*
	Progesterone	2.011 (1.057-3.826) *p=0.032*	1.965 (1.115-3.460) *p=0.019*	2.465 (1.260-4.876) *p=0.009*
	LH	1.859 (1.005-3.427) *p=0.046*	1.922 (1.106-3.335) *p=0.020*	2.511 (1.294-4.992) *p=0.007*
	FSH	1.909 (1.030-3.524) *p=0.038*	2.003 (1.148-3.491) *p=0.014*	2.524 (1.302-4.993) *p=0.006*
	ER-α SB	1.898 (1.013-3.539) *p=0.043*	1.851 (1.059-3.223) *p=0.029*	2.371 (1.212-4.673) *p=0.012*
	ER-β SB	1.810 (0.972-3.350) p=0.059	1.882 (1.077-3.275) *p=0.025*	2.378 (1.216-4.685) *p=0.011*
	AR SB	1.724 (0.917-3.210) p=0.086	1.804 (1.027-3.152) *p=0.038*	2.357 (1.206-4.644) *p=0.012*

OR values for top-bottom classification are based on controls only. *OR - for single hormones not adjusted in relation to other hormones or SB, only age adjusted. **OR - for hormones adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with italic and after bonferonni correction p-values ≤ 0.0007 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Table 4-27: Joint association of androgens (top quintiles) with risk of breast cancer – all cases.

Hormones		Joint association of androgens	
		Androstenedione	
		Testosterone	DHEAS
		OR* (95% CI)	**OR* (95% CI)**
Not adjusted		2.567 (1.703-4.678) **p<0.0001**	1.972 (1.106-3.483) *p=0.015*
		OR (95% CI)**	**OR** (95% CI)**
Adjusted	Oestradiol	2.703 (1.663-4.417) **p<0.0001**	2.047 (1.170-3.576) *p=0.012*
	Oestrone	2.759 (1.718-4.457) **p<0.0001**	2.052 (1.175-3.577) *p=0.011*
	Androstenedione		
	Testosterone		1.213 (0.632-2.319) p=0.559
	DHEAS	3.354 (1.945-5.903) **p<0.0001**	
	SHBG	2.458 (1.538-3.941) **p<0.0001**	1.995 (1.193-3.337) *p=0.008*
	Progesterone	3.827 (2.168-6.940) **p<0.0001**	2.586 (1.288-5.306) *p=0.008*
	LH	2.665 (1.667-4.281) **p<0.0001**	2.125 (1.218-3.704) *p=0.008*
	FSH	2.577 (1.617-4.121) **p<0.0001**	1.982 (1.143-3.425) *p=0.014*
	ER-α SB	2.603 (1.620-4.198) **p<0.0001**	2.039 (1.175-3.531) *p=0.003*
	ER-β SB	2.586 (1.614-4.158) **p<0.0001**	2.052 (1.180-3.560) *p=0.003*
	AR SB	2.491 (1.553-4.010) **p<0.0001**	1.885 (1.081-3.270) *p=0.004*

OR values for top-bottom classification are based on controls only. *OR - for single hormones not adjusted in relation to other hormones or SB, only age adjusted. **OR - for hormones adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with italic and after bonferonni correction p-values ≤ 0.0007 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Table 4-28: Joint association of androgens and progesterone (top quintiles) with risk of breast cancer – all cases.

Hormones		Joint association of androgens and progesterone	
		Progesterone	
		Androstenedione	Testosterone
		OR* (95% CI)	**OR* (95% CI)**
Not adjusted		1.731 (0.989-2.866) *p=0.043*	1.824 (1.092-2.972) *p=0.019*
		OR (95% CI)**	**OR** (95% CI)**
Adjusted	Oestradiol	1.741 (0.532-2.996) *p=0.046*	1.847 (1.104-3.080) *p=0.019*
	Oestrone	1.740 (1.006-2.960) *p=0.041*	2.016 (1.209-3.357) *p=0.007*
	Androstenedione		1.330 (0.731-2.404) p=0.346
	Testosterone	1.031 (1.018-2.146) p=0.924	
	DHEAS	2.175 (0.547-2.575) *p=0.022*	2.635 (1.345-5.315) *p=0.005*
	SHBG	1.711 (1.184-2.926) *p=0.050*	1.802 (1.082-2.992) *p=0.023*
	Progesterone		
	LH	1.770 (0.995-3.024) *p=0.037*	1.883 (1.132-3.123) *p=0.014*
	FSH	1.708 (1.123-2.912) *p=0.050*	1.820 (1.095-3.015) *p=0.020*
	ER-α SB	1.678 (0.981-2.848) p=0.056	1.806 (1.082-3.003) *p=0.023*
	ER-β SB	1.661 (0.965-2.839) p=0.064	1.808 (1.084-3.004) *p=0.022*
	AR SB	1.653 (0.959-2.829) p=0.068	1.863 (1.112-3.111) *p=0.017*

OR values for top-bottom classification are based on controls only. *OR - for single hormones not adjusted in relation to other hormones or SB, only age adjusted. **OR - for hormones adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Table 4-29: Joint association of androgens and gonadotrophins (top quintiles) with risk of breast cancer – all cases.

Hormones		Joint association of androgens and gonadotrophins		
		Androstenedione	Testosterone	
		LH	LH	FSH
		OR* (95% CI)	**OR* (95% CI)**	**OR* (95% CI)**
Not adjusted		1.731 (0.935-4.174) *p=0.033*	3.029 (1.284-5.555) *p=0.003*	5.924 (2.337-16.152) **p<0.0001**
		OR (95% CI)**	**OR** (95% CI)**	**OR** (95% CI)**
Adjusted	Oestradiol	1.741 (1.099-5.154) *p=0.027*	2.964 (1.445-6.222) *p=0.003*	6.261 (2.535-17.712) **p<0.0001**
	Oestrone	1.740 (1.253-6.111) *p=0.012*	3.532 (1.672-7.748) *p=0.001*	7.518 (2.893-23.355) **p<0.0001**
	Androstenedione		2.033 (0.932-4.472) p=0.074	4.350 (1.671-12.703) **p=0.004**
	Testosterone	1.031 (0.790-3.931) p=0.164		
	DHEAS	2.175 (1.069-5.147) *p=0.032*	3.277 (1.542-7.150) *p=0.002*	5.979 (2.385-17.077) **p<0.0001**
	SHBG	1.711 (1.011-4.675) *p=0.045*	2.987 (1.447-6.318) *p=0.003*	6.351 (2.509-18.325) **p<0.0001**
	Progesterone	2.664 (1.084-5.207) **p=0.030**	2.879 (1.360-6.224) *p=0.006*	4.978 (1.047-14.347) **p<0.0001**
	LH			9.117 (3.425-27.654) **p<0.0001**
	FSH	1.708 (1.066-4.939) *p=0.032*	3.153 (1.523-6.689) *p=0.002*	
	ER-α SB	2.455 (1.137-5.339) *p=0.022*	3.085 (1.502-6.488) *p=0.002*	5.816 (2.296-16.703) **p<0.0001**
	ER-β SB	2.272 (1.062-4.866) *p=0.033*	3.026 (1.476-6.353) *p=0.002*	5.495 (2.179-15.718) **p<0.0001**
	AR SB	2.293 (1.072-4.914) *p=0.031*	3.058 (1.490-6.424) *p=0.002*	5.594 (2.216-16.017) **p<0.0001**

OR values for top-bottom classification are based on controls only. *OR - for single hormones not adjusted in relation to other hormones or SB, only age adjusted. **OR - for hormones adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with italic and after bonferonni correction p-values ≤ 0.0007 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Table 4-30: Joint association of oestrogens (top quintiles) with risk of breast cancer – cases that gave a sample less than 2 years before diagnosis.

Hormones		Joint association of oestrogens Oestradiol Oestrone
		OR* (95% CI)
Not adjusted		2.578 (1.044-6.120) *p=0.033*
		OR (95% CI)**
Adjusted	Oestradiol	
	Oestrone	
	Androstenedione	2.413 (0.962-5.770) p=0.051
	Testosterone	2.101 (0.830-5.061) p=0.103
	DHEAS	2.562 (0.885-6.058) *p=0.035*
	Progesterone	2.512 (1.013-5.942) *p=0.039*
	SHBG	2.402 (0.963-5.727) *p=0.051*
	LH	2.532 (1.021-5.991) *p=0.037*
	FSH	2.622 (1.057-6.209) *p=0.031*
	ER-α SB	2.283 (0.885-5.529) p=0.073
	ER-β SB	2.283 (0.883-5.537) p=0.074
	AR SB	2.280 (0.883-5.521) p=0.074

OR values for top-bottom classification are based on controls only. *OR - for single hormones not adjusted in relation to other hormones or SB, only age adjusted. **OR - for hormones adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with italic and after bonferonni correction p-values ≤ 0.0007 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Table 4-31: Joint association of oestrogens and androgens (top quintiles) with risk of breast cancer – cases that gave a sample less than 2 years before diagnosis.

Hormones		Joint association of oestrogens and androgens	
		Oestrone	
		Androstenedione	Testosterone
		OR* (95% CI)	**OR* (95% CI)**
Not adjusted		2.496 (1.255-6.725) *p=0.029*	3.390 (1.818-8.384) *p=0.001*
		OR (95% CI)**	**OR** (95% CI)**
Adjusted	Oestradiol	2.390 (0.972-5.571) *p=0.048*	3.719 (1.710-8.017) *p=0.001*
	Oestrone		
	Androstenedione		2.883 (1.276-6.388) *p=0.009*
	Testosterone	1.792 (0.698-4.377) p=0.208	
	DHEAS	2.510 (1.045-5.778) *p=0.033*	3.478 (1.593-7.511) *p=0.001*
	Progesterone	2.730 (1.100-6.569) *p=0.026*	3.275 (1.509-7.009) *p=0.002*
	SHBG	2.307 (0.976-5.206) *p=0.048*	2.996 (1.394-6.344) *p=0.004*
	LH	2.406 (1.023-5.401) *p=0.037*	3.393 (1.585-7.149) *p=0.001*
	FSH	2.494 (1.063-5.585) *p=0.029*	3.438 (1.605-2.427) *p=0.001*
	ER-α SB	2.229 (0.918-5.094) p=0.064	3.128 (1.435-6.668) *p=0.003*
	ER-β SB	2.396 (0.979-5.541) *p=0.045*	3.274 (1.495-7.028) *p=0.002*
	AR SB	2.229 (0.917-5.099) p=0.064	3.135 (1.435-6.700) *p=0.003*

OR values for top-bottom classification are based on controls only. *OR - for single hormones not adjusted in relation to other hormones or SB, only age adjusted. **OR - for hormones adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with italic and after bonferonni correction p-values ≤ 0.0007 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Table 4-32: Joint association of androgens (top quintiles) with risk of breast cancer – cases that gave a sample less than 2 years before diagnosis.

Hormones		Joint association of androgens	
		Androstenedione	
		Testosterone	DHEAS
		OR* (95% CI)	**OR* (95% CI)**
Not adjusted		2.555 (1.426-4.580) *p=0.001*	2.046 (0.959-3.655) *p=0.038*
		OR (95% CI)**	**OR** (95% CI)**
Adjusted	Oestradiol	2.818 (1.531-5.132) **p<0.0001**	2.211 (1.086-4.362) *p=0.024*
	Oestrone	2.689 (1.479-4.818) *p=0.001*	2.051 (1.006-4.047) *p=0.042*
	Androstenedione		
	Testosterone		1.349 (0.597-2.993) p=0.465
	DHEAS	3.090 (1.581-6.058) *p=0.001*	
	SHBG	2.360 (1.308-4.194) *p=0.004*	1.948 (0.962-1.052) p=0.056
	Progesterone	3.346 (1.658-6.822) *p=0.001*	2.324 (0.978-5.593) p=0.056
	LH	2.607 (1.449-4.620) *p=0.001*	2.187 (1.079-4.297) *p=0.025*
	FSH	2.572 (1.430-4.554) *p=0.001*	2.064 (1.024-4.022) *p=0.037*
	ER-α SB	2.614 (1.433-4.692) *p=0.001*	2.120 (1.052-4.133) *p=0.030*
	ER-β SB	2.619 (1.437-4.699) *p=0.001*	2.203 (1.087-4.326) *p=0.024*
	AR SB	2.490 (1.374-4.435) *p=0.002*	2.049 (1.018-3.895) *p=0.038*

OR values for top-bottom classification are based on controls only. *OR - for single hormones not adjusted in relation to other hormones or SB, only age adjusted. **OR - for hormones adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with italic and after bonferonni correction p-values ≤ 0.0007 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Table 4-33: Joint association of androgens and progesterone (top quintiles) with risk of breast cancer – cases that gave a sample less than 2 years before diagnosis.

Hormones		Joint association of androgens and progesterone	
		Progesterone	
		Androstenedione	Testosterone
		OR* (95% CI)	**OR* (95% CI)**
Not adjusted		1.928 (0.978-3.590) *p=0.048*	2.025 (1.094-3.705) *p=0.023*
		OR (95% CI)**	**OR** (95% CI)**
Adjusted	Oestradiol	2.046 (1.030-3.943) *p=0.035*	2.137 (1.131-3.945) **p=0.017**
	Oestrone	1.789 (0.907-3.408) p=0.083	2.121 (1.122-3.913) *p=0.018*
	Androstenedione		1.560 (0.739-3.223) p=0.235
	Testosterone	1.239 (0.568-2.641) p=0.583	
	DHEAS	2.340 (1.045-5.251) *p=0.038*	2.883 (1.263-6.788) *p=0.013*
	SHBG	1.868 (0.947-3.565) p=0.063	1.971 (1.047-3.618) *p=0.031*
	Progesterone		
	LH	1.971 (1.004-3.743) *p=0.042*	2.078 (1.109-3.801) *p=0.019*
	FSH	1.893 (0.965-3.591) p=0.056	2.030 (1.082-3.715) *p=0.024*
	ER-α SB	1.728 (0.872-3.292) p=0.104	1.952 (1.031-3.593) *p=0.035*
	ER-β SB	1.819 (0.914-3.486) p=0.078	1.975 (1.042-3.640) *p=0.032*
	AR SB	1.782 (0.897-3.404) p=0.088	2.011 (1.060-3.711) *p=0.028*

OR values for top-bottom classification are based on controls only. *OR - for single hormones not adjusted in relation to other hormones or SB, only age adjusted. **OR - for hormones adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with italic and after bonferonni correction p-values ≤ 0.0007 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Table 4-34: Joint association of androgens and gonadotrophins (top quintiles) with risk of breast cancer – cases that gave a sample less than 2 years before diagnosis.

Hormones		Joint association of androgens and gonadotrophins	
		Testosterone	
		LH	FSH
		OR* (95% CI)	**OR* (95% CI)**
Not adjusted		3.816 (1.527-7.891) *p=0.001*	6.404 (2.620-21.648) *p=0.001*
		OR (95% CI)**	**OR** (95% CI)**
Adjusted	Oestradiol	3.799 (1.649-8.663) *p=0.001*	7.068 (2.546-21.342) **p<0.0001**
	Oestrone	4.386 (1.846-10.462) **p=0.0001**	8.322 (2.844-27.583) **p<0.0001**
	Androstenedione	2.332 (0.899-5.793) p=0.072	4.137 (1.301-13.522) *p=0.015*
	Testosterone		
	DHEAS	4.231 (1.737-10.337) *p=0.001*	6.406 (2.224-19.798) **p<0.0001**
	SHBG	3.690 (1.590-8.501) *p=0.002*	6.496 (2.231-20.388) **p<0.0001**
	Progesterone	3.792 (1.581-9.057) *p=0.003*	5.368 (1.807-16.804) *p=0.003*
	LH		9.227 (2.981-31.158) **p<0.0001**
	FSH	3.916 (1.686-9.014) *p=0.001*	
	ER-α SB	3.828 (1.664-8.709) *p=0.001*	5.633 (1.901-17.607) *p=0.002*
	ER-β SB	3.866 (1.677-8.823) *p=0.001*	5.756 (1.941-18.012) *p=0.002*
	AR SB	3.843 (1.669-8.759) *p=0.001*	5.675 (1.916-17.736) *p=0.002*

OR values for top-bottom classification are based on controls only. *OR - for single hormones not adjusted in relation to other hormones or SB, only age adjusted. **OR - for hormones adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR OR with p-values ≤0.05 marked with italic and after bonferonni correction p-values ≤ 0.0007 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Table 4-35: Joint association of oestrogens and androgens (top quintiles) with risk of breast cancer – cases that gave a sample more than 2 years before diagnosis.

Hormones		Joint association of oestrogens and androgens	
		Oestradiol	
		Androstenedione	Testosterone
		OR* (95% CI)	**OR* (95% CI)**
Not adjusted		2.101 (1.114-5.069) *p=0.047*	2.362 (1.250-4.786) *p=0.010*
		OR (95% CI)**	**OR** (95% CI)**
Adjusted	Oestradiol		
	Oestrone	2.079 (0.968-4.261) *p=0.050*	2.328 (1.185-4.438) *p=0.012*
	Androstenedione		1.907 (0.924-3.827) p=0.073
	Testosterone	1.328 (0.573-2.959) p=0.496	
	DHEAS	2.201 (1.004-4.624) *p=0.041*	2.559 (1.256-5.056) *p=0.007*
	SHBG	2.089 (0.971-4.291) *p=0.050*	2.388 (1.213-4.566) *p=0.010*
	Progesterone	2.424 (1.087-5.222) *p=0.026*	2.541 (1.265-4.981) *p=0.007*
	LH	2.027 (0.942-4.162) *p=0.060*	2.296 (1.166-4.383) *p=0.013*
	FSH	2.108 (0.978-4.343) *p=0.048*	2.388 (1.207-4.589) *p=0.010*
	ER-α SB	2.240 (1.034-4.646) *p=0.033*	2.307 (1.169-4.414) *p=0.013*
	ER-β SB	2.101 (0.976-4.315) *p=0.048*	2.344 (1.191-4.476) *p=0.011*
	AR SB	1.927 (0.875-4.011) p=0.088	2.211 (1.108-4.257) *p=0.020*

OR values for top-bottom classification are based on controls only. *OR - for single hormones not adjusted in relation to other hormones or SB, only age adjusted. **OR - for hormones adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Table 4-36: Joint association of androgens (top quintiles) with risk of breast cancer – cases that gave a sample more than 2 years before diagnosis.

Hormones		Joint association of androgens Androstenedione Testosterone
		OR* (95% CI)
Not adjusted		2.558 (1.450-4.666) *p=0.001*
		OR (95% CI)**
Adjusted	Oestradiol	2.564 (1.396-4.648) *p=0.002*
	Oestrone	2.772 (1.535-4.935) **p<0.0001**
	Androstenedione	
	Testosterone	
	DHEAS	3.494 (1.766-6.979) **p<0.0001**
	SHBG	2.505 (1.394-4.424) *p=0.002*
	Progesterone	4.236 (2.090-8.776) **p<0.0001**
	LH	2.569 (1.430-4.542) *p=0.001*
	FSH	2.564 (1.428-4.529) *p=0.001*
	ER-α SB	2.606 (1.441-4.461) *p=0.001*
	ER-β SB	2.556 (1.419-4.530) *p=0.001*
	AR SB	2.467 (1.362-4.389) *p=0.002*

OR values for top-bottom classification are based on controls only. *OR - for single hormones not adjusted in relation to other hormones or SB, only age adjusted. **OR - for hormones adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR OR with p-values ≤0.05 marked with italic and after bonferonni correction p-values ≤ 0.0007 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Table 4-37: Joint association of androgens and gonadotrophins (top quintiles) with risk of breast cancer – cases that gave a sample more than 2 years before diagnosis.

Hormones		Joint association of androgens and gonadotrophins
		Testosterone
		FSH
		OR* (95% CI)
Not adjusted		5.330 (1.456-15.174) *p=0.004*
		OR (95% CI)**
Adjusted	Oestradiol	5.325 (1.7071-17.167) *p=0.004*
	Oestrone	6.499 (1.998-22.751) *p=0.002*
	Androstenedione	4.273 (1.331-14.095) *p=0.014*
	Testosterone	
	DHEAS	5.556 (1.759-18.166) *p=0.003*
	SHBG	5.670 (1.802-18.468) *p=0.003*
	Progesterone	4.628 (1.390-15.495) *p=0.011*
	LH	7.304 (2.228-25.269) *p=0.001*
	FSH	
	ER-α SB	6.064 (1.962-19.731) *p=0.002*
	ER-β SB	5.023 (1.604-16.213) *p=0.005*
	AR SB	5.490 (1.759-17.698) *p=0.003*

OR values for top-bottom classification are based on controls only. *OR - for single hormones not adjusted in relation to other hormones or SB, only age adjusted. **OR - for hormones adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with italic and after bonferonni correction p-values ≤ 0.0007 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Joint associations of SB between sex steroid receptors were also investigated. Women who gave a sample >2 years before diagnosis with ER-α and ER-β SB or ER-α and AR SB or ER-β and AR SB in the top quintile had 1.950 (95% CI: 1.072-3.963; P≤0.05) or 1.981 (95% CI: 1.093-3.779; P≤0.05) or 2.482 (95% CI: 1.072-3.963; P≤0.05) fold breast cancer risk. These joint associations though lost significance after bonferonni correction (Table 4-38). Statistically significant association of the joint effect of steroid receptor SB was not seen when all cases and these who gave a sample less than two years before breast cancer diagnosis were investigated (Appendix VII).

Further analysis was carried out to investigate joint association of each receptor's SB with each hormone. When all cases were analysed, women with ER-α and testosterone in top quintiles had 1.999 (95% CI: 1.224-4.836; p≤0.05) fold breast cancer which did not remain statistically significant after adjustment with androstenedione (Table 4-39). Women with AR SB and oestrone or testosterone in the top quintiles had 2.480 (95% CI: 1.177-5.179; p≤0.05) and 2.558 (95% CI: 1.389-6.163; p≤0.05) fold breast cancer respectively (Table 4-40 and 4-41). Joint association between AR SB and androstenedione also showed to increase breast cancer (OR: 2.410; 95% CI: 1.124-5.170; p≤0.05) but did not remain significant after adjustment with testosterone (Tables 4-41). All the above joint associations with breast cancer risk were lost though after bonferonni correction. Any association between SB of the sex steroid receptors and progesterone and gonadotrophins did not show significant association with breast cancer risk (Appendix VIII).

For those women who had given a sample ≤2 years before diagnosis joint association of AR SB and oestrone referred to 2.575 (95%CI: 1.054-6.168; p≤0.05) fold breast cancer risk (Table 4-42). All other pair combinations were shown to be statistically insignificant associated with breast cancer risk (Appendix VIII). For those women who had given a sample >2 years before diagnosis, women with ER-α or ER-β SB in the top quintile along with testosterone were at a 2.304 (95%CI: 1.330-7.031; p≤0.05) and 2.754 (95%CI: 1.436-7.313; p≤0.05) fold breast cancer risk which did not remain statistically significant after adjustment with androstenedione respectively (Table 4-43 and 4-44). Women with ER-β SB and SHBG in the top quintile were at 7.306 (95%CI: 1.436-7.313; p≤0.05) fold breast cancer risk which remained after all adjustments (Table 4-45). Women with AR SB and testosterone in top quintile had 2.807 (95%CI: 1.321-7.882; p≤0.05) fold breast cancer risk which remained statistically significant after adjustments apart for androstenedione (Table 4-46). In addition, joint association of AR SB and androstenedione referred to 2.961 (95%CI: 1.124-7.311; p≤0.05) fold breast cancer which did not remain statistically significant after adjustment with testosterone (Table 4-46). Significance of all the above joint association with breast cancer risk though was lost after bonferonni correction. All the other joint associations were not shown to be statistically significant associated with breast cancer risk (Appendix VIII).

Table 4-38: Joint association of high steroid receptor serum bioactivity (top quintiles) with risk of breast cancer - cases that gave a sample more than 2 years before diagnosis.

Joint association of high SB of steroid receptors – more than 2 years before diagnosis				
SB		**ER-α and ER-β**	**ER-α and AR**	**ER-β and AR**
		OR* (95% CI)	**OR* (95% CI)**	**OR* (95% CI)**
Not adjusted		1.950 (1.072-3.963) *p=0.031*	1.981 (1.093-3.779) *p=0.036*	2.482 (1.072-3.963) *p=0.007*
		OR (95% CI)**	**OR** (95% CI)**	**OR** (95% CI)**
Adjusted	Oestradiol	1.987 (1.059-3.452) *p=0.028*	1.924 (0.990-3.604) *p=0.046*	2.445 (1.241-4.676) *p=0.008*
	Oestrone	1.897 (1.010-3.483) *p=0.040*	1.991 (1.024-3.738) *p=0.036*	2.407 (1.217-4.622) *p=0.009*
	Androstenedione	1.897 (1.010-3.483) *p=0.050*	1.932 (0.990-3.640) *p=0.046*	2.305 (1.163-4.429) *p=0.014*
	Testosterone	1.904 (1.009-3.358) *p=0.040*	1.872 (0.957-3.528) p=0.058	2.428 (1.218-4.703) *p=0.009*
	DHEAS	1.950 (1.042-3.403) *p=0.031*	1.991 (1.027-3.724) *p=0.035*	2.524 (1.278-4.845) *p=0.006*
	SHBG	1.844 (0.982-3.513) *p=0.050*	1.956 (1.006-3.670) *p=0.041*	2.384 (1.207-4.571) *p=0.010*
	Progesterone	1.935 (1.033-3.513) *p=0.034*	1.944 (1.003-3.635) *p=0.042*	2.532 (1.282-4.861) *p=0.006*
	LH	1.917 (1.022-3.486) *p=0.037*	1.901 (0.979-3.563) *p=0.050*	2.401 (1.217-4.599) *p=0.009*
	FSH	1.938 (1.034-3.523) *p=0.033*	1.970 (1.014-3.694) *p=0.039*	2.472 (1.255-4.727) *p=0.007*
	ER-α SB			1.801 (0.836-3.802) p=0.126
	ER-β SB		1.735 (0.825-3.551) p=0.137	
	AR SB	1.891 (0.951-3.552) p=0.060		

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with italic and after bonferonni correction p-values ≤ 0.0007 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Table 4-39: Joint association of high ER-α serum bioactivity and androgens (top quintiles) with risk of breast cancer - all cases.

SB/Hormones		Joint association of SB and androgens
		ER-α SB
		Testosterone
		OR* (95% CI)
Not adjusted		1.996 (1.224-4.836) *p=0.044*
		OR (95% CI)**
Adjusted	Oestradiol	2.077 (1.039-4.159) *p=0.037*
	Oestrone	1.951 (0.988-3.845) p=0.052
	Androstenedione	1.582 (0.776-3.207) p=0.202
	Testosterone	
	DHEAS	1.997 (1.004-3.966) *p=0.047*
	SHBG	1.975 (0.995-3.914) *p=0.050*
	Progesterone	1.981 (0.995-3.935) *p=0.050*
	LH	1.967 (0.996-3.878) *p=0.049*
	FSH	1.990 (1.005-3.930) *p=0.046*
	ER-α SB	
	ER-β SB	2.031 (1.010-4.978) *p=0.045*
	AR SB	1.989 (0.986-4.004) p=0.053

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with italic and after bonferonni correction p-values ≤ 0.0007 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Table 4-40: Joint association of high AR serum bioactivity and oestrogens (top quintiles) with risk of breast cancer - all cases.

SB - Hormones		Joint association of SB and oestrogens AR SB Oestrone
		OR* (95% CI)
Not adjusted		2.480 (1.177-5.179) *p=0.015*
		OR (95% CI)**
Adjusted	Oestradiol	2.305 (0.917-4.149) *p=0.027*
	Oestrone	
	Androstenedione	2.167 (0.804-3.722) *p=0.013*
	Testosterone	2.540 (0.899-3.244) *p=0.026*
	DHEAS	2.488 (0.938-3.634) *p=0.015*
	SHBG	2.454 (1.173-5.221) *p=0.017*
	Progesterone	2.626 (1.192-5.636) p=0.093
	LH	2.384 (1.083-5.031) p=0.102
	FSH	2.518 (1.215-5.314) p=0.072
	ER-α SB	2.294 (1.083-4.931) *p=0.030*
	ER-β SB	2.586 (1.215-5.598) *p=0.014*
	AR SB	

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with italic and after bonferonni correction p-values ≤ 0.0007 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Table 4-41: Joint association of high AR serum bioactivity and androgens (top quintiles) with risk of breast cancer - all cases.

SB - Hormones		Joint association of SB and androgens	
		AR SB	
		Androstenedione	Testosterone
		OR* (95% CI)	**OR* (95% CI)**
Not adjusted		2.410 (1.124-5.170) *p=0.024*	2.558 (1.389-6.163) *p=0.012*
		OR (95% CI)**	**OR** (95% CI)**
Adjusted	Oestradiol	2.380 (0.790-5.197) *p=0.027*	2.554 (0.985-5.246) *p=0.013*
	Oestrone	2.356 (1.093-5.146) *p=0.029*	2.440 (1.173-5.156) *p=0.017*
	Androstenedione		2.104 (0.686-4.550) *p=0.050*
	Testosterone	1.709 (1.068-3.847) p=0.189	
	DHEAS	2.466 (1.110-5.459) *p=0.023*	2.663 (1.225-5.279) *p=0.011*
	SHBG	2.254 (1.042-4.939) *p=0.039*	2.531 (1.210-5.385) *p=0.014*
	Progesterone	2.333 (1.130-5.146) *p=0.033*	2.599 (1.260-5.584) *p=0.013*
	LH	2.336 (0.997-5.106) *p=0.030*	2.501 (1.099-5.280) *p=0.014*
	FSH	2.398 (1.145-5.246) *p=0.026*	2.548 (1.244-5.387) *p=0.012*
	ER-α SB	2.207 (0.997-4.941) *p=0.050*	2.347 (1.098-5.086) *p=0.028*
	ER-β SB	2.548 (1.115-5.757) *p=0.020*	2.650 (1.244-5.741) *p=0.012*
	AR SB		

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with italic and after bonferonni correction p-values ≤ 0.0007 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Table 4-42: Joint association of high AR serum bioctivity and oestrogens (top quintiles) with risk of breast cancer – cases that gave a sample less than 2 years before diagnosis.

SB - Hormones		Joint association of SB and oestrogens
		AR SB
		Oestrone
		OR* (95% CI)
Not adjusted		2.575 (1.054-6.168) *p=0.033*
		OR (95% CI)**
Adjusted	Oestradiol	2.587 (1.045-6.109) *p=0.033*
	Oestrone	
	Androstenedione	2.550 (1.025-6.050) *p=0.036*
	Testosterone	2.460 (0.969-5.984) *p=0.050*
	DHEAS	2.536 (1.025-5.985) *p=0.036*
	SHBG	2.408 (0.968-5.723) *p=0.050*
	Progesterone	2.724 (1.092-6.523) *p=0.026*
	LH	2.522 (1.020-5.951) *p=0.037*
	FSH	2.626 (1.061-6.202) *p=0.030*
	ER-α SB	2.761 (1.088-6.733) *p=0.027*
	ER-β SB	3.308 (1.285-8.283) *p=0.011*
	AR SB	

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with italic and after bonferonni correction p-values ≤ 0.0007 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Table 4-43: Joint association of high ER-α serum bioctivity and andoregens (top quintiles) with risk of breast cancer – cases that gave a sample more than 2 years before diagnosis.

SB - Hormones		Joint association of SB and androgens
		ER-α SB
		Testosterone
		OR* (95% CI)
Not adjusted		2.304 (1.329-7.031) *p=0.042*
		OR (95% CI)**
Adjusted	Oestradiol	2.288 (0.971-5.124) *p=0.049*
	Oestrone	2.270 (0.978-4.995) *p=0.046*
	Androstenedione	1.776 (0.736-4.604) p=0.183
	Testosterone	
	DHEAS	2.375 (1.008-5.326) *p=0.039*
	SHBG	2.313 (0.993-5.110) *p=0.043*
	Progesterone	2.413 (1.023-5.424) *p=0.036*
	LH	2.226 (0.956-4.912) p=0.053
	FSH	2.288 (0.980-5.068) *p=0.046*
	ER-α SB	
	ER-β SB	2.046 (0.862-4.606) p=0.091
	AR SB	2.234 (0.939-5.054) p=0.059

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with italic and after bonferonni correction p-values ≤ 0.0007 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Table 4-44: Joint association of high ER-β serum bioactivity and androgens (top quintiles) with risk of breast cancer - cases that gave a sample more than 2 years before diagnosis.

SB - Hormones		Joint association of SB and androgens ER-β SB Testosterone
		OR* (95% CI)
Not adjusted		2.754 (1.436-7.313) *p=0.012*
		OR (95% CI)**
Adjusted	Oestradiol	2.768 (1.202-6.151) *p=0.014*
	Oestrone	2.570 (1.093-5.769) *p=0.025*
	Androstenedione	2.032 (0.827-4.771) p=0.109
	Testosterone	
	DHEAS	2.864 (1.241-6.388) *p=0.011*
	SHBG	2.553 (1.118-5.607) *p=0.021*
	Progesterone	2.886 (1.248-6.459) *p=0.011*
	LH	2.763 (1.210-6.078) *p=0.013*
	FSH	2.743 (1.206-6.002) *p=0.013*
	ER-α SB	2.332 (1.002-5.213) *p=0.042*
	ER-β SB	
	AR SB	2.859 (1.226-6.45) *p=0.012*

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with italic and after bonferonni correction p-values ≤ 0.0007 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Table 4-45: Joint association of high ER-β serum bioactivity and sex hormone bidning globulin (top quintiles) with risk of breast cancer - cases that gave a sample more than 2 years before diagnosis.

SB/Hormones		Joint association of SB and SHBG ER-β SB SHBG
		OR* (95% CI)
Not adjusted		7.306 (1.769-36.412) *p=0.007*
		OR (95% CI)**
Adjusted	Oestradiol	7.090 (1.702-35.184) *p=0.008*
	Oestrone	7.156 (1.716-35.562) *p=0.008*
	Androstenedione	8.039 (1.901-40.424) *p=0.005*
	Testosterone	8.493 (1.997-42.868) *p=0.004*
	DHEAS	7.309 (1.754-36.301) *p=0.007*
	SHBG	
	Progesterone	7.290 (1.744-36.337) *p=0.008*
	LH	7.071 (1.688-35.268) *p=0.009*
	FSH	7.212 (1.722-35.968) *p=0.008*
	ER-α SB	6.084 (1.423-30.770) *p=0.017*
	ER-β SB	
	AR SB	6.824 (1.606-34.383) *p=0.010*

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Table 4-46: Joint association of high AR serum bioactivity and androgens (top quintiles) with risk of breast cancer - cases that gave a sample more than 2 years before diagnosis.

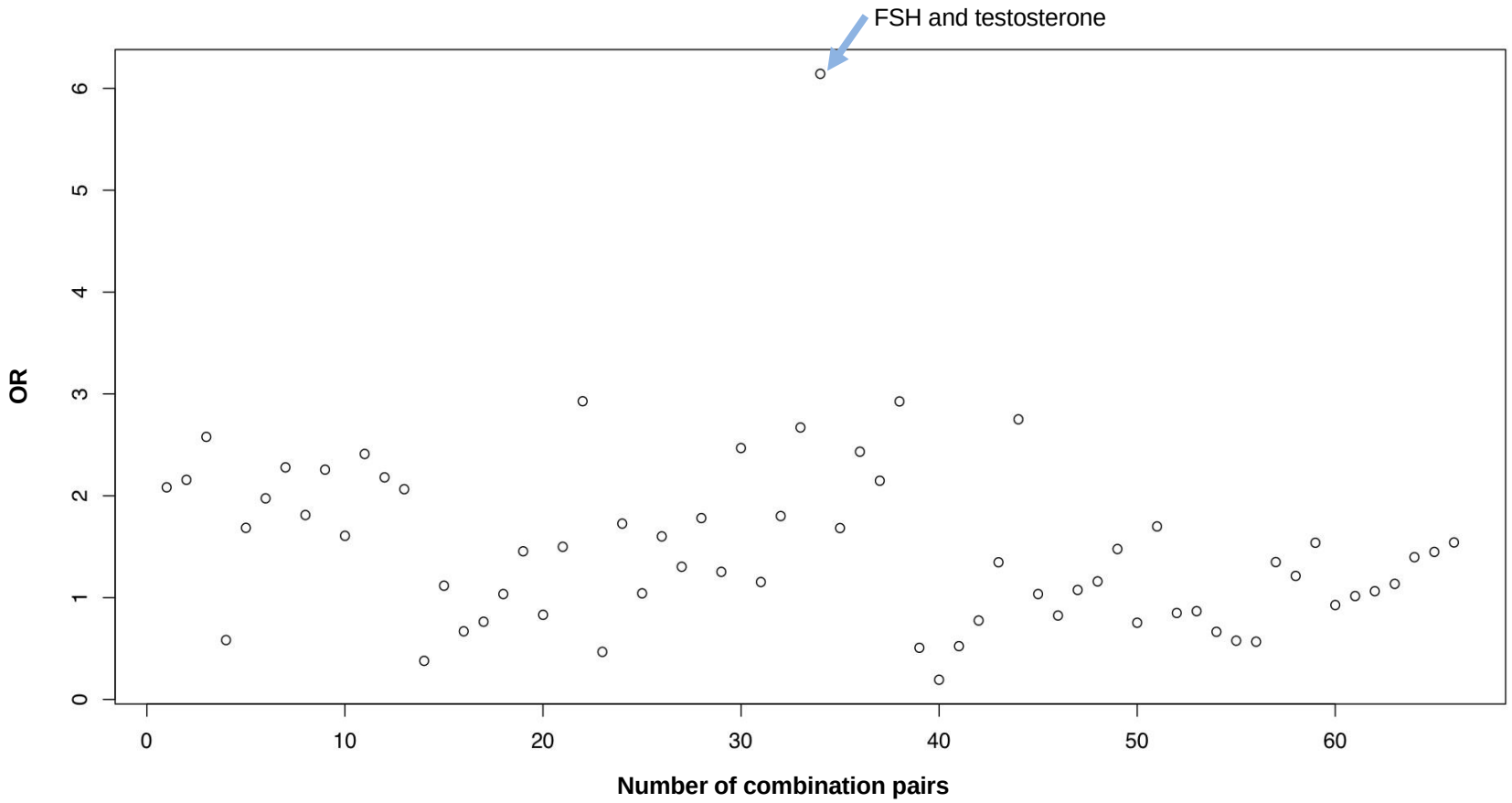
SB - Hormones		Joint association of SB and androgens	
		AR SB	
		Androstenedione	Testosterone
		OR* (95% CI)	**OR* (95% CI)**
		2.961 (1.234-7.310) *p=0.016*	2.807 (01.321-7.882) *p=0.020*
		OR (95% CI)**	**OR** (95% CI)**
Adjusted	Oestradiol	2.872 (1.147-6.901) *p=0.020*	2.693 (1.081-6.404) *p=0.027*
	Oestrone	2.949 (1.180-7.083) *p=0.017*	2.766 (1.117-6.538) *p=0.022*
	Androstenedione		2.229 (0.865-5.487) p=0.086
	Testosterone	2.122 (0.813-5.311) p=0.112	
	DHEAS	3.180 (1.247-7.841) *p=0.012*	3.174 (1.235-7.892) *p=0.013*
	SHBG	2.939 (1.127-6.789) *p=0.022*	2.830 (1.136-6.739) *p=0.020*
	Progesterone	2.939 (1.151-7.237) *p=0.020*	3.141 (1.222-7.805) *p=0.014*
	LH	2.900 (1.157-6.980) *p=0.019*	2.719 (1.094-6.448) *p=0.025*
	FSH	2.941 (1.173-7.083) *p=0.017*	2.786 (1.120-6.619) *p=0.022*
	ER-α SB	2.371 (0.917-5.885) p=0.066	2.213 (0.860-5.434) p=0.088
	ER-β SB	2.600 (1.011-6.427) *p=0.040*	2.452 (0.963-5.959) p=0.051
	AR SB		

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with italic and after bonferonni correction p-values ≤ 0.0007 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

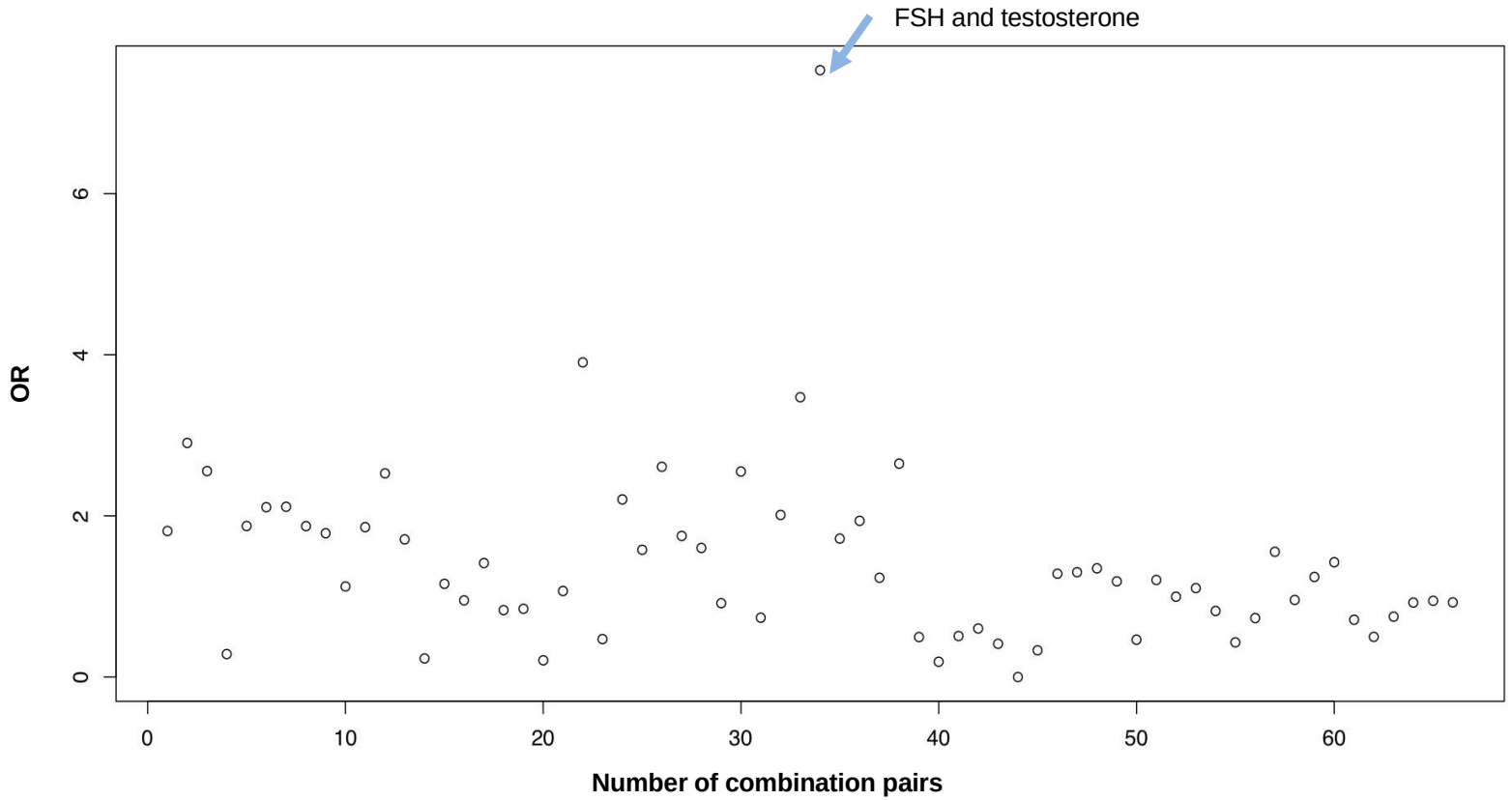
The following figures (Figure 4-1, 4-2, 4-3) summarise all different pairs investigated indicating FSH and testosterone to have the best predictive power in comparison to all other pairs examined independently of time to diagnosis. ER-β and SHBG pair also has high breast cancer predictive power but this time only more than two years before diagnosis.

Figure 4-1: Odds ratio of all possible combination pairs of hormones/serum bioactivity- all cases.



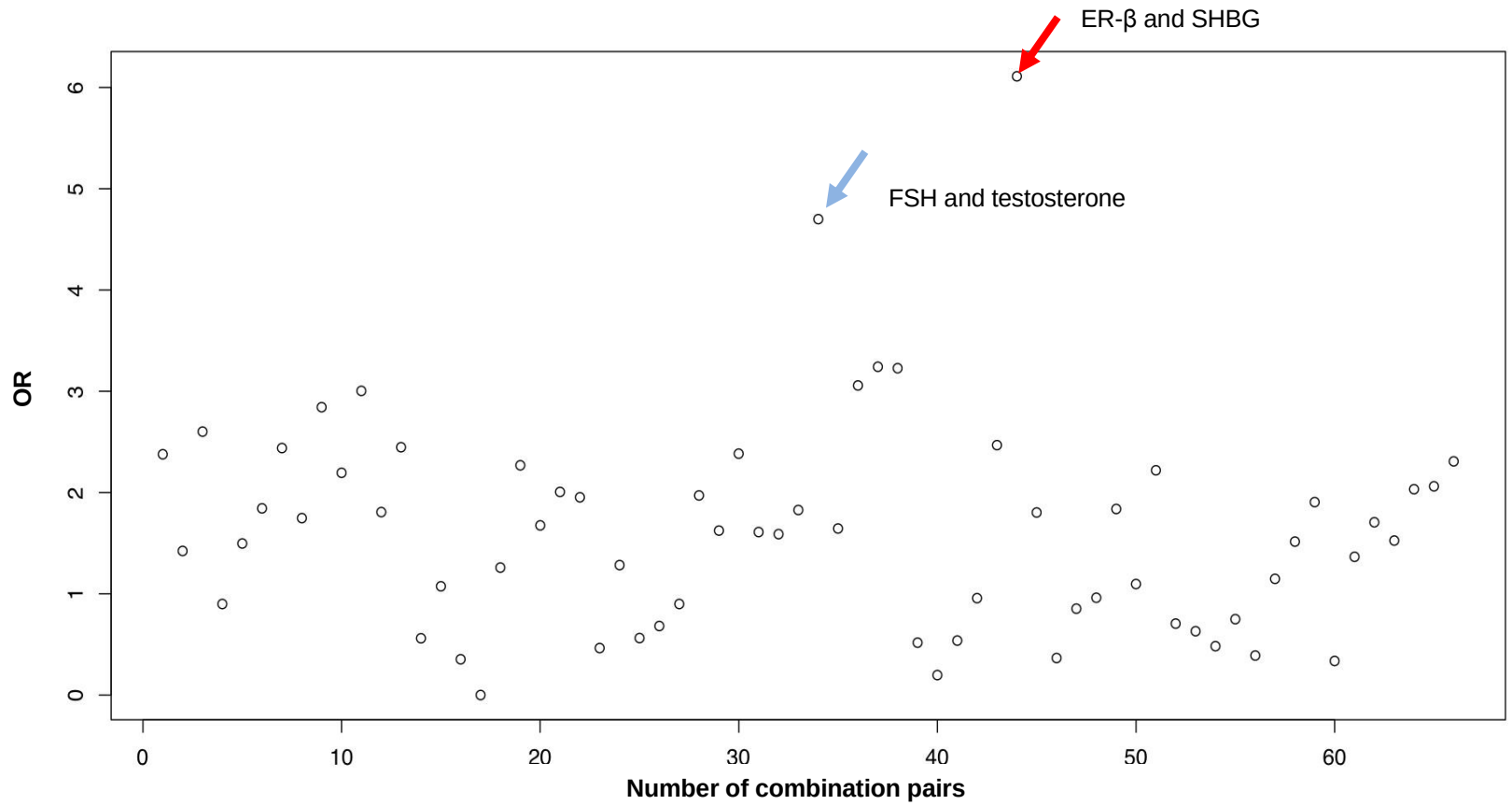
FSH=follicle stimulating hormone; OR=odds ratio; SB=serum bioactivity

Figure 4-2: Odds ratio of all possible combination pairs of hormones/serum bioactivity- cases that gave a sample less than two years before breast cancer diangnosis.



FSH=follicle stimulating hormone; OR=odds ratio; SB=serum bioactivity

Figure 4-3: Odds ratio of all possible combination pairs of hormones/serum bioactivity- cases that gave a sample more than two years before breast cancer diangnosis.



ER=oestrogen receptor; FSH=follicle stimulating hormone; OR=odds ratio; SB=serum bioactivity, SHBG=sex hormones binding globulin

4.3.7 Validation of the predictive power of testosterone and FSH in breast cancer risk

To further validate the predictive power of testosterone and FSH, 1000 experiments were run, where 10% of the data was removed from cases and 10% of controls and OR was re-evaluated. The quintiles were also re-calculated after removing 10% of the data. Distribution (Gaussian distribution) of the OR values obtained with a mean of 5.637; p-value=0.002, median of 5.443; p-value=0.001 and variance of 1.824; p-value=2.77E-05.

4.3.8 Examination of the synergistic effect of the different pairs investigated

From the analyses undertaken in section 4.3.6, possible synergistic interplays between sex steroids, gonadotrophins and sex steroid receptor SB among postmenopausal women were found to be associated with increased breast cancer. Further investigation, was carried out by computing the observed versus expected OR ratio to quantify whether a hidden synergistic effect is possible. Initially the eexpected OR was estimated under the assumption that two hormones are independent (Table 4-47). Then based on the observed and the expected OR the ratio was calculated (Table 4-48). As indicated in Table 4-48, a significant hidden synergistic effect was confirmed for FSH and testosterone (p=0.048) and for ER-β and SHBG (p=0.022).

Table 4-47: Expected odds ratio of the different combinations of hormones/SB to investigate their synergistic effect in breast cancer.

Hormone / SB	Hormone / SB	Expected OR	Lower CI	Upper CI	p-value
Oestradiol	Oestrone	1.47	0.87	2.58	0.152
Oestradiol	Androstenedione	1.69	0.97	3.05	0.062
Oestradiol	Testosterone	1.92	1.10	3.51	**0.021**
Oestradiol	DHEAS	1.09	0.61	1.89	0.782
Oestradiol	Progesterone	1.17	0.66	2.05	0.590
Oestradiol	SHBG	0.53	0.26	1.03	0.062
Oestradiol	LH	0.75	0.40	1.31	0.315
Oestradiol	FSH	1.06	0.58	1.83	0.855
Oestradiol	ER-α	1.38	0.84	2.46	0.205
Oestradiol	ER-β	1.11	0.63	1.96	0.715
Oestradiol	AR	1.20	0.72	2.15	0.493
Oestrone	Androstenedione	2.19	1.40	3.80	**0.001**
Oestrone	Testosterone	2.50	1.57	4.36	**0.000**
Oestrone	DHEAS	1.55	0.79	2.38	0.204
Oestrone	SHBG	0.65	0.33	1.31	0.233
Oestrone	Progesterone	1.46	0.86	2.57	0.158
Oestrone	LH	0.94	0.51	1.66	0.827
Oestrone	FSH	1.35	0.76	2.29	0.308
Oestrone	ER-α	1.76	1.20	3.06	**0.004**
Oestrone	ER-β	1.39	0.82	2.45	0.222
Oestrone	AR	1.52	0.95	2.69	0.079
Androstenedione	Testosterone	2.97	1.80	5.18	**0.000**
Androstenedione	DHEAS	2.25	0.88	2.83	0.090
Androstenedione	Progesterone	1.68	0.97	3.04	0.065
Androstenedione	SHBG	0.76	0.37	1.57	0.451
Androstenedione	LH	1.09	0.57	2.01	0.799
Androstenedione	FSH	1.59	0.85	2.73	0.150
Androstenedione	ER-α	2.07	1.35	3.64	**0.001**
Androstenedione	ER-β	1.61	0.92	2.93	0.096
Androstenedione	AR	1.78	1.09	3.19	**0.022**
Testosterone	DHEAS	2.95	0.99	3.25	0.053
Testosterone	Progesterone	1.90	1.09	3.49	**0.023**
Testosterone	SHBG	0.86	0.42	1.81	0.691
Testosterone	LH	1.24	0.63	2.33	0.529
Testosterone	FSH	1.82	0.95	3.15	0.072
Testosterone	ER-α	2.36	1.51	4.16	**0.000**
Testosterone	ER-β	1.83	1.03	3.36	**0.039**
Testosterone	AR	2.04	1.22	3.66	**0.006**

Hormone / SB	Hormone / SB	Expected OR	Lower CI	Upper CI	p-value
DHEAS	SHBG	0.38	0.24	0.94	**0.037**
DHEAS	Progesterone	1.08	0.60	1.88	0.796
DHEAS	LH	0.59	0.36	1.21	0.149
DHEAS	FSH	0.95	0.53	1.69	0.859
DHEAS	ER-α	1.46	0.76	2.26	0.252
DHEAS	ER-β	1.03	0.58	1.81	0.916
DHEAS	AR	1.18	0.65	1.98	0.587
SHBG	Progesterone	0.53	0.26	1.03	0.060
SHBG	LH	0.33	0.16	0.61	**0.000**
SHBG	FSH	0.46	0.23	0.90	**0.023**
SHBG	ER-α	0.61	0.32	1.25	0.178
SHBG	ER-β	0.50	0.25	0.98	**0.043**
SHBG	AR	0.53	0.28	1.09	0.083
Progesterone	LH	0.75	0.39	1.30	0.310
Progesterone	FSH	1.06	0.58	1.82	0.858
Progesterone	ER-α	1.38	0.83	2.44	0.211
Progesterone	ER-β	1.11	0.63	1.95	0.716
Progesterone	AR	1.19	0.71	2.14	0.505
LH	FSH	0.66	0.35	1.16	0.147
LH	ER-α	0.88	0.49	1.58	0.674
LH	ER-β	0.71	0.38	1.25	0.239
LH	AR	0.76	0.43	1.37	0.361
FSH	ER-α	1.27	0.73	2.19	0.396
FSH	ER-β	1.00	0.56	1.75	0.989
FSH	AR	1.09	0.63	1.91	0.765
ER-α	ER-β	1.31	0.79	2.35	0.288
ER-α	AR	1.43	0.91	2.56	0.117
ER-β	AR	1.14	0.68	2.04	0.625

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicle stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Table 4-48: Observed versus expected odds ratio of the different combinations of hormones/SB to quantify their possible synergistic effect in breast cancer.

Hormone / SB	Hormone / SB	Observed/Expected OR	Lower CI	Upper CI	p-value
Oestradiol	Oestrone	1.35	0.51	3.37	0.541
Androstenedione	Oestradiol	1.12	0.47	2.47	0.799
Oestradiol	Testosterone	1.01	0.44	2.16	0.973
Oestradiol	DHEAS	0.98	0.39	2.52	0.971
Oestradiol	Progesterone	0.94	0.38	2.33	0.887
Oestradiol	SHBG	0.77	0.18	3.48	0.738
Oestradiol	LH	0.94	0.26	3.67	0.924
Oestradiol	FSH	0.69	0.16	3.18	0.631
Oestradiol	ER-α	1.01	0.38	2.45	0.983
Oestradiol	ER-β	0.67	0.23	1.95	0.466
Oestradiol	AR	1.28	0.48	3.09	0.619
Androstenedione	Oestrone	0.81	0.33	1.89	0.628
Oestrone	Testosterone	0.96	0.40	2.15	0.923
Oestrone	DHEAS	0.77	0.33	2.24	0.634
Oestrone	SHBG	0.72	0.19	2.73	0.631
Oestrone	Progesterone	1.09	0.42	2.68	0.865
Oestrone	LH	1.11	0.37	3.47	0.852
Oestrone	FSH	1.09	0.39	3.20	0.865
Oestrone	ER-α	0.87	0.35	1.96	0.743
Oestrone	ER-β	0.83	0.33	1.99	0.671
Oestrone	AR	1.64	0.63	3.77	0.311
Androstenedione	Testosterone	0.83	0.40	1.64	0.588
Androstenedione	DHEAS	0.80	0.52	2.49	0.704
Androstenedione	Progesterone	0.96	0.43	2.01	0.920
Androstenedione	SHBG	0.73	0.19	2.85	0.655
Androstenedione	LH	1.89	0.72	5.12	0.198
Androstenedione	FSH	1.28	0.46	3.76	0.633
Androstenedione	ER-α	0.88	0.35	1.98	0.756
Androstenedione	ER-β	0.98	0.37	2.40	0.964
Androstenedione	AR	1.35	0.50	3.24	0.555
Testosterone	DHEAS	0.52	0.39	1.82	0.306
Testosterone	Progesterone	0.91	0.41	1.88	0.808
Testosterone	SHBG	1.25	0.35	4.39	0.733
Testosterone	LH	2.29	0.88	6.12	0.091
Testosterone	FSH	3.08	1.01	9.98	**0.048**
Testosterone	ER-α	0.82	0.34	1.82	0.630
Testosterone	ER-β	0.99	0.38	2.38	0.986
Testosterone	AR	1.24	0.47	2.94	0.658

Hormone / SB	Hormone / SB	Observed/Expected OR	Lower CI	Upper CI	p-value
DHEAS	SHBG	2.36	0.52	6.85	0.265
DHEAS	Progesterone	1.12	0.53	2.40	0.766
DHEAS	LH	1.60	0.55	3.70	0.392
DHEAS	FSH	1.47	0.46	4.71	0.515
DHEAS	ER-α	0.63	0.25	1.96	0.424
DHEAS	ER-β	1.02	0.36	3.01	0.971
DHEAS	AR	0.95	0.35	2.68	0.922
SHBG	Progesterone	1.03	0.28	4.08	0.965
SHBG	LH	0.68	0.13	3.78	0.660
SHBG	FSH	1.26	0.40	4.09	0.692
SHBG	ER-α	2.42	0.65	8.62	0.186
SHBG	ER-β	6.67	1.31	35.49	**0.022**
SHBG	AR	2.26	0.62	7.58	0.215
Progesterone	LH	1.15	0.44	3.32	0.771
Progesterone	FSH	1.02	0.30	3.80	0.970
Progesterone	ER-α	0.99	0.34	2.66	0.978
Progesterone	ER-β	0.65	0.21	2.00	0.450
Progesterone	AR	1.42	0.50	3.66	0.510
LH	FSH	1.25	0.57	3.00	0.576
LH	ER-α	0.73	0.24	2.24	0.582
LH	ER-β	0.85	0.27	2.93	0.796
LH	AR	0.77	0.18	3.33	0.731
FSH	ER-α	0.91	0.30	2.75	0.865
FSH	ER-β	1.63	0.57	4.98	0.363
FSH	AR	0.84	0.26	2.58	0.759
ER-α	ER-β	1.06	0.48	2.14	0.882
ER-α	AR	0.99	0.44	1.92	0.983
ER-β	AR	1.52	0.66	3.22	0.330

AR=androgen receptor; CI=confidence interval; DHEAS=dehydro- epiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

4.4 Discussion

The role of sex steroid hormones in breast cancer has been known for a long time and has been the subject of many studies. In this study, we report for the first time on sex steroid hormone bioactivity. Serum ER-α and ER-β bioactivity using a yeast-based assay was significantly higher in postmenopausal women prior to diagnosis of invasive ER-positive breast cancer compared to controls. Women with ER-α SB in the top quintile more than two 2 years before diagnosis had a two-fold breast cancer risk increase. We further validated the role of sex steroid hormones in breast cancer risk. Less than 2 years before diagnosis, oestrone was associated with increased breast cancer risk. Testosterone and androstenedione levels were shown to be significantly associated with increased cancer risk irrespective of time (>6months and <5 years) before invasive ER-positive breast cancer, with the first being independent of other hormones and the latter being dependent on testosterone.

For the first time joint associations of sex steroid hormones, gonadotrophins and sex steroid receptor SB were examined. Joint association of FSH and testosterone was shown to be highly associated with breast cancer risk with women with both hormones in top quintiles having almost six fold increased breast cancer risk independent of time to diagnosis with further analysis confirming a possible synergistic effect. Interestingly, SHBG and ER-β were also shown to be associated with high breast cancer risk with women with both in the top quintiles more than two years before diagnosis having a six fold increased breast cancer risk

The strengths of this study are (1) use of standardised protocol for serum sample collection with protocol adherence confirmed by the lack of any difference in mean hormone or steroid receptor SB levels between the different trial centres (data not shown) (2) prospective nested case-control design which ensured the selection controls from the same population (trial participants) as that in which the breast cancer cases occurred (3) confirmation of breast cancer diagnosis and receptor status from the treating physicians which eliminated possible misidentification of cases from use of cancer registry data or self reporting alone (4) use of women not on HRT with ER positive invasive breast cancer which ensured that a homogenous case mix (5) measurement of all sex steroid hormones of ovarian origin with only DHEA and oestrone sulphate not being analysed (6) analysis of the joint association of hormones, hormones and SB and SBs provided robust results. Ideally cases should have been a random selection from all women with fully characterised ER positive breast cancer within the trial. However the need to start experimental work meant that the first 200 fully characterized cases that fulfilled eligibility criteria were used.

Our findings are in keeping with our previous findings of elevated ER-α and ER-β SB in women with breast cancer at the time of clinical diagnosis [162] and with the meta-analysis by Key *et al*. that showed sex steroid hormone levels more than 2 years prior to diagnosis to be more significantly associated with breast cancer risk [197]. Free oestradiol has the highest known affinity for ER-α and ER-β receptors [340] and a statistically significant correlation between free oestradiol and receptor SB was found in our study. Phosphorylation of the receptors is probably modulated by other surrogates as well. Thus, we found increased breast cancer risk in women with ER-α SB in the highest quintile more than two years before

diagnosis in the absence of a correlation with individual oestrogens. In our previous study, receptor SB was also 2-3 folds higher than the actual oestradiol concentration [162]. Other factors such as insulin growth factor-1 (IGF-1) which have been shown to bind to the ER [341] could contribute to the higher bioactivity. IGF-1 has been shown to stimulate ER-mediated trans-activation and ER-phosphorylation [341]. Recently elevated IGF-1 levels have been associated with ER-positive breast-cancer risk [342]. We hope to investigate IGF-1 levels in relation to ER SB within our study cohort where appropriate samples (spun within 24 hours [343]) are available. Additionally, other serum steroid independent co-activators may have an impact on breast carcinogenesis through ER-α and ER-β SB activation, such as cAMP and cytokines [344]. Generally, the advantage of using SB assays for steroid receptors is that their levels reflect the sum of all the factors in the serum that trans-activate the two different ERs. Given the significant reduction in breast cancer incidence in women taking anti-oestrogens such as tamoxifen [126], raloxifene [127] and aromatase inhibitors [128], it is likely that ER SB may prove to be beneficial in individualising and monitoring breast cancer chemopreventive strategies. Studies are urgently needed to assess this further.

The role of ER-α in breast carcinogenesis has been extensively studied and it is one of the main tumour markers used in the clinical setting. However, the role of ER-β still remains to be determined. Cell-based assays have shown ER-β to be less active on gene transcription than ER-α [345]. This could be the explanation for our findings of that while ER-β SB is different among cases and controls more than 2 years before diagnosis, levels in the top quintile are not associated with an increased breast cancer risk. Expression of both receptors favours a positive response to endocrine therapy [345] but it is unclear whether the addition of ER-β to

ER-α as a tumour marker would be clinically beneficial. Recently, our group showed that women whose levels of ER-α and ER-β SB in the top quintiles at the time of diagnosis had a 10 fold increased risk for ER positive breast cancer [162]. In this study, joint association of SB of ERs was shown to be associated with breast cancer risk in those women who gave a sample more than two years before diagnosis indicating that SB of sex steroid receptors could prove useful for breast cancer risk assessment. Women with SB in the top quintiles had a breast cancer risk ranging between 2-2.5 fold. This reinforces the possibility that SB may prove to be a useful tool in individualising and monitoring breast cancer chemopreventive strategies.

Oestrone is the main circulating oestrogen after menopause in postmenopausal women [346]. We found oestrone in the top quintile less than two years before breast cancer diagnosis to be associated with increased risk. Our findings support previous studies which showed only oestrone and not oestradiol or bio-available oestradiol to be significantly associated with increased breast cancer risk [212, 218, 224, 225]. Studies that have not been able to find a significant association of oestrone with breast cancer risk are on the whole based on small number of cases [203, 217, 347]. There are conflicting results in the literature on the role of oestradiol. There have been three reports in postmenopausal women that similar to our study did not find any differences between oestradiol levels in cases and controls [226, 347] but there are conflicting reports as well [212, 217, 222-225] that demonstrate an increased risk [203, 218, 219]. While earlier publications from the European Prospective Investigation into Cancer and Nutrition (EPIC) study [198] found total oestradiol, oestradiol and oestrone to be associated with breast cancer risk independent of time to diagnosis, a recent study by Zeleniuch-

Jacquotte *et al* found that oestrone close to diagnosis had the strongest association with increased breast cancer risk [199].

In general, some of these differences observed in studies between the hormones and their association with breast cancer may be a result of the variety of assays used to measure oestrogens, direct or indirect radioimmunoassay (RIA) and immunoassays. The reliability and validity of steroid sex hormone measurements in biologic specimens using immunoassays has been recently evaluated. While considerable variation was found in results from different laboratories, the measurements from a single laboratory was reproducible [151]. The most sensitive method is mass spectrometry which does not lend itself to use in clinical settings [150] [149]. Immunoassays have been shown that to yield similar estimates of most sex steroid hormones in comparison to mass spectrometry [151, 152]. We have in addition used calculated free oestradiol and testosterone levels as these were highly correlated with the actual hormone levels measured on equilibrium dialysis [348].

Oestrone levels did not correlate significantly with ER SB, raising the question of how this oestrogen might exert its effect on breast carcinogenesis. Oestrone is a weak oestrogen which preferentially binds to the alpha receptor but with low affinity [349]. Studies have shown that oestrogens also exert their effects through their binding to the oestrogen G protein-coupled receptor GPR30 (GPER), which is independent of ER-α and ER-β [350]. This has led to the suggestion that oestrone may be capable of inducing ERK phosphorylation via GPR30 without the requirement of ER receptors, as oestradiol has been shown previously [350]. In addition, other oestrogen metabolites independent of ER mediation has been

shown to contribute to breast carcinogenesis [351]. If this is confirmed, then there could be implications for hormonal therapy in prevention and treatment of breast cancer in postmenopausal women.

Androstenedione and testosterone were associated with a two-fold increase in breast cancer risk independent of time from diagnosis. Overall, there is conflicting data on endogenous levels of androgens and breast cancer risk [201]. In the meta-analysis high testosterone levels were associated with breast cancer risk in postmenopausal women [197] with six out of nine studies included in the pooled study showing an increased risk of breast cancer in women having testosterone in the highest quintiles [218, 219, 222-225]. This observation was also seen in the EPIC study [198], confirming no difference in the effect of testosterone in relation to time of diagnosis. Our observation that androstenedione was significantly associated with breast cancer risk is also in line with most of the previous prospective studies [203, 218, 224, 225]. But in contrast to other studies [203, 218, 221, 224, 225], our results showed DHEAS to be associated with breast cancer risk only after adjustment with androstenedione. Since, both DHEAS and androstenedione are largely of adrenal origin in postmenopausal women, our data suggests that adrenal androgens may play a role in breast carcinogenesis.

The role of endogenous androgens in breast cancer development has been debated. One of the possible pathways is through increased aromatase activity in the setting of oestrogen depletion after menopause and increased capacity to convert testosterone to oestradiol and androstenedione to oestrone may be the major factor. After adjustment for oestradiol and oestrone levels, the association of the androgens with breast cancer risk was shown to remain in the main,

indicating that androgens may have oestrogen-independent effects on breast cancer, an observation that has been reported by other authors [197, 198, 202]. It is possible that androgens may influence breast cancer risk by directly binding to AR to stimulate or inhibit breast cell growth [352, 353] but we were unable to demonstrate any association between AR SB and breast cancer. Joint association of AR though with testosterone when all cases were analysed, demonstrated that women with AR SB and testosterone in top quintiles have a 2.5 fold increased risk which remained significant after all adjustments. Additionally, joint association of AR SB and oestrone in samples taken less than two years before breast cancer diagnosis, showed that women with AR SB and oestrone in the top quintiles had a 2.5 fold increased risk which remained significant after all adjustments. A second possible explanation for the direct association of androgens with breast cancer risk is conversion of high circulatory levels of these hormones to oestrogen either locally in the breast or peripherally in adipose tissue with the oestrogens then being responsible for tumour development [198]. A statistically significant correlation between free testosterone and both ER receptors was observed. While free testosterone is the best ligand of AR, androgens have also been shown to bind and activate ERs [353], favouring the view that a third pathway may exist through binding to ER and directly promoting breast cell proliferation.

Joint association of high levels of oestrone and testosterone was shown to increase breast cancer risk up to 3.5 folds independent of time to diagnosis. During the last few years exogenous oestrogens and testosterone have been extensively used to manage post-menopausal symptoms. Recently, it was demonstrated by the Nurses' Health Study that women using combination of

oestrogen and testosterone therapy rather than oestrogen alone had a higher breast cancer risk [354]. In addition, in a recent trial in women undergoing adjuvant treatment for breast cancer, use of tibolone, a drug which is a synthetic steroid with oestrogenic, progestational and androgenic properties to prevent side-effects of vasomotor symptoms and bone loss was shown to be associated with breast cancer recurrence [355]. The increasing number of studies showing association of oestrogens and testosterone with breast cancer risk including the results of the current study suggests that use of such combinations in postmenopausal women should be carefully considered.

Progesterone has been shown in *in-vitro* studies and animal work to both decrease and increase breast cancer risk [165]. Controversy surrounds the true effect of progesterone on breast proliferation with various progestins being shown to block, stimulate or have no effect on cell growth [356]. The only reported large study investigating endogenous levels of progesterone did not show any association of the circulating progesterone with breast cancer risk [202] in agreement with these data. Epidemiological studies, however, have consistently shown an increase in breast cancer risk when exogenous progesterone is used in combination with oestradiol [177, 179] causing decline in HRT use among postmenopausal women [357]. Joint association of progesterone with testosterone or androstenedione was shown to increase breast cancer risk by almost 2 fold but this significance was lost after adjustments. Combination of high endogenous levels of progesterone and oestrogens was not shown to be associated with increased risk of breast cancer although, exogenous administration of oestrogen in combination with progestin is responsible for a higher breast cancer risk than oestrogen alone [179].

Our results on breast cancer risk reduction with increasing levels of SHBG are similar to meta-analysis [197] with only one study demonstrating no reduction in risk with high levels of SHBG [358]. Conflicting data has been shown with regards to SHBG levels in samples that were collected at different time intervals in relation to time of diagnosis. Similar to the meta-analysis elevated SHBG levels less than two years before diagnosis were shown to have a stronger association with reduced breast cancer risk [197]. This was in contrast to Kaaks *et al* who reported no difference [198]. Given that SHBG binds to both oestradiol and testosterone prohibiting binding of the hormones to their receptors, it is expected that high levels of SHBG are associated with lower risk.

Interestingly though, joint association of SHBG with ER-β was shown to increase breast cancer risk by six fold more than two years before diagnosis. Previous studies have demonstrated that high SHBG levels were associated with higher mammographic density, indicating a positive relationship between them [359, 360] with mammographic density shown to be a strong independent predisposing factor for breast cancer [62]. SHBG has also been suggested to promote the effects of oestradiol by interacting with plasma membrane binding sites in target cells within the breast [359]. ER-β is the main ER expressed in epithelial cells in normal human breast. Therefore, it could be hypothesised that the observed increased breast cancer risk in our study associated with high levels of SHBG and ER-β SB may be due to high breast proliferation caused through ER-β possibly promoted by SHBG which also associates with high breast density. Further studies though are needed before any suggestions can be made.

There are not that many studies investigating LH and FSH levels and breast cancer risk. The only study identified was by Wang *et al*, who reported in 1976 that breast cancer patients with low LH and FSH levels measured at the point of diagnosis had a faster recurrence, even though the data did not reach statistical significance [361]. In this study, it was shown for the first time that high level of LH was associated with a decreased breast cancer risk after adjustment for androstenedione. On the other hand, FSH was not shown to be association with breast cancer.

Joint association of the gonadotrophins or in combination with oestrogens were not shown to be associated with breast cancer risk, but a significantly increased risk was observed for FSH and testosterone up to 5 years prior to breast cancer diagnosis, with the risk ranging 5-6 folds. This is an interesting finding as it is for the first time shown androgens in combination with gonadotrophins to be associated with such an increased breast cancer risk. Validation of these observations is required in independent studies. It is difficult to explain in postmenopausal women with inactive ovaries since the known action of FSH is in the ovaries where it stimulates granulosa cells to produce oestradiol through aromatisation of testosterone [362] and increases production of androgens in response to LH by stimulating secretion of paracrine factors, such as inhibin B [363]. One explanation may be related to the action of FSH on the breast. FSHR mRNA is expressed in the normal breast and has been also detected in breast cancer cell lines and breast cancer core biopsies [364]. Therefore, an attractive model in postmenopausal women would be that FSH through its binding to FSHR in the breast tissue modulates aromatase activity triggering the conversion of androgens, in particular testosterone, into oestrogens that in turn are responsible

for breast proliferation. Additionally, it could also be speculated that FSH could be a surrogate marker of IGF1 since previous studies have shown a positive correlation between the two hormones [360]. Previous work on bovine granulosa cells demonstrated that IGF1 with FSH act synergistically to up-regulate sex steroid synthesis [365]. Animal work has indicated a synergism of sex steroid hormones and FSH in bovine granulosa cells. DHT was shown to synergise with FSH inducing FSHR [366]. Based on our hypothesis, we have set up collaboration with Professor Louis Dubeau to further investigate the possible synergistic effect of FSH and testosterone in breast carcinogenesis in a mouse model. Two transgenic mouse lines expressing Cre recombinase under the control of a truncated form of the FSHR will be crossed with a ROSA26R Cre reporter mouse [367]. Expression of FSHR in the breast will be examined (stroma, epithelium and fat). If FSHR mRNA is confirmed then mouse mammary cells will be treated with FSH and IGF1 alone and in combination to investigate whether increased aromatization occurs by measuring CYP11A1, HSD3B1 and CYP19A1 (genes known to encode aromatase enzymes) mRNA and oestradiol, oestrone levels and compare to non treated cells. Finally, the relationship of FSH levels and HRT use would be interesting. HRT is known to increase breast cancer risk and possibly decrease FSH levels. In this study women with high levels of FSH and testosterone have increased risk to develop breast cancer – therefore, it could be hypothesised that an altered pathway could be responsible for the development of breast cancer between the action of FSH and sex steroid hormones under the administration of HRT.

Sex steroid hormones and their receptors along with gonadotrophins could prove a useful tool for the early detection and prevention of breast cancer. Currently,

even though mammography screening has been suggested to decrease mortality from breast cancer [117], its true value has been the subject of much debate [118]. Tools such as the Gail, Claus and Golditz models using family history and lifestyle data have been developed to calculate a woman's absolute risk of breast cancer and triage women to screening or risk reducing mastectomy. When oestradiol was added in the Golditz model the test's concordance statistics were slightly increased [123]. It is possible that addition of oestrone, testosterone and FSH, or the bioactivity of ER receptors may improve the performance of these scores. The significance of the attempts to increase the models performance is in identification of those women at high risk, eventually leading to improved overall survival rates of breast cancer patients through prevention strategies.

In summary, in postmenopausal women with ER-positive breast cancer, testosterone and androstenedione independent of time to diagnosis, ER-α SB more than 2 years prior to diagnosis and oestrone levels less than 2 years before diagnosis were associated with increased risk. Sex steroid receptor SB assays may be a new tool for breast cancer risk assessment and warrant further research. Combination of the different hormones/SB of the sex steroid hormone receptors has been shown to have a better prediction power in relation to single hormones or SB. Testosterone with FSH independent of time to diagnosis and SHBG with ER-β more than two years before breast cancer diagnosis were shown to be highly significant associated with breast cancer risk indicating a synergistic effect in mammary carcinogenesis. Understanding the complex signals that hormones convey to the mammary gland could shed light on the events that surround breast tumour formation and growth and eventually initiate new strategies for treatment.

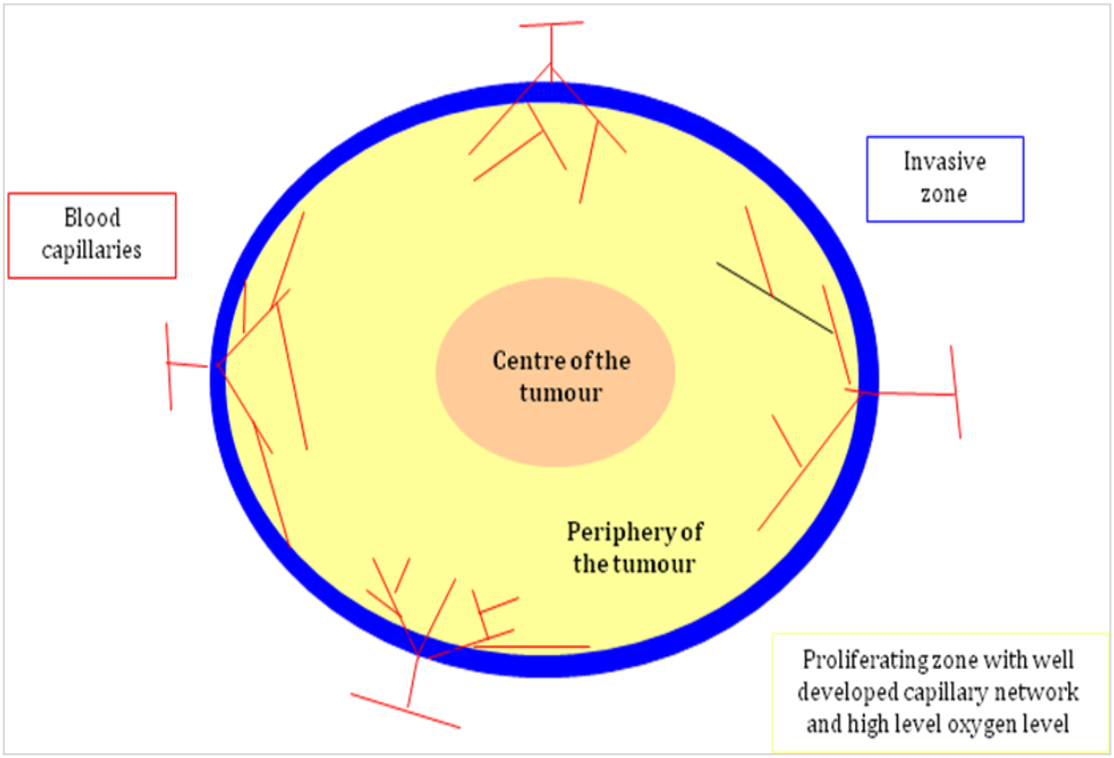
5 DNA METHYLATION IN BREAST CANCER

5.1 Introduction

Identification of novel molecular markers with the potential for optimal breast cancer management and improved survival rates is essential. During the last decade a huge emphasis has been given to the identification of genetic changes and expression profiles that correlate with clinical characteristics of the disease, in an attempt to discover genetic markers predicting prognosis and response to treatment. Many of these studies have been based on a single sample from within the tumour, assuming that this single region reflects the genetic signature of the whole cancer. However, there are increasing reports of the presence of intra-tumour heterogeneity and its effect on expression profiling in several cancer types [368-374], including breast cancer [375, 376].

Intra-tumour heterogeneity is the result of a multi-factorial microenvironment which exhibits a zonal heterogeneity from central to peripheral regions [377] (Figure 5-1). Studies comparing the central with the peripheral zone have identified expression of different molecules within these regions. In the centre, which is characterised by hypoxic conditions, genes such as vascular endothelial growth factor (VEGF) have been shown to be regulated and molecules such as the matrix metalloproteinases (MMPs) are over-expressed. In the tumour periphery, which is localised at the stromal border, forming the biologically active invasion front and cancer stem cell reservoir, molecules such as E-cadherin have been shown to be down-regulated [378, 379].

Figure 5-1: Intra-tumour heterogeneity. Intra-tumoural heterogeneity is a result of a multi-factorial microenvironment which exhibits a zonal separation from central to peripheral regions.



Even, for ER and PR, which are the most significant markers in breast cancer treatment strategies, there is conflicting data with studies also reporting to be differently expressed in different regions within the tumour [380-382]. This has led to an increased emphasis for the need to study more than one part of the tumour to ensure the generation of accurate and reproducible data, especially as these data are used to guide patient management. Based on these observations it was hypothesised that the problem of intra-tumour heterogeneity may be overcome by studying DNA based alterations, such as epigenetic changes, specifically DNA methylation, which may not be affected by the zonal microenvironment of the tumour.

In addition, to improve early diagnostic strategies better models are needed to investigate early stage disease. Studies have indicated that epigenetic alterations may be the initiating events in the expansion of cells in pre-neoplastic lesions [271]. However, although epigenetic alterations contribute to the pathogenesis of breast cancer, the effect of these changes as the initiating event of carcinogenesis has been difficult to study. Our group was amongst the first to describe an epigenetic field defect, altered DNA methylation in morphological normal breast tissue adjacent to tumour, in breast cancer [383, 384] and these findings have been confirmed by others [273, 385, 386]. If methylation changes arise early in normal tissues, leading to regional epigenetic defects, then a comparison between histological normal tissues from cancer patients and healthy controls could lead to the identification of methylation markers that could be useful in risk assessment [387]. Since most of the research studies to identify markers for the early detection of the disease have been focused on cells within the tumour from early stage cases we have hypothesised that by studying morphologically normal tissue adjacent to the tumour it could prove a good strategy for the identification of risk prediction markers. Such study is important as the identification of patients who are at high risk of cancer could benefit from prophylactic treatment.

Multiple genes have been shown to be differently methylated in normal versus tumour tissue (discussed in section 2.3.4). Recently, our group and others demonstrated that PCGT genes are more likely to have cancer specific promoter DNA hypermethylation than non-PCGT genes [279, 280, 388]. Furthermore, we showed that hypermethylation of *NEUROD1*, a PCGT gene, within pretreatment core biopsies preferentially discriminated between neoplastic and non-neoplastic breast tissue samples and was associated with a favourable response to

treatment [305]. Given the data showing that expression profiles are affected by intra-tumour heterogeneity, it is essential to establish that the methylated profile of the core biopsy is representative of the entire tumour. PCGT methylation could prove a good candidate marker for serum analysis but before analysing serum it is important to investigate whether such cancer specific PCGT methylation is representative of the entire tumour. In addition, it is interesting to investigate whether PCGT methylation in morphologically normal tissue adjacent to the tumour could predict the presence of breast cancer.

For the purposes of this study; **1)** initially the expression and DNA methylation profile of *NEUROD1* in 9 breast cancer cell lines and 63 frozen breast cancer tissues were analysed. Based on the results it was postulated that DNA methylation signature may carry information independently from the expression profile of the tumour. In order to further investigate both **2)** the predictive role and **3)** homogeneity of PCGT gene methylation in breast cancer, methylation analysis was carried out using MethyLight: a highly sensitive real-time PCR methylation assay [250]. 55 PCGT genes (6 of which are also known ER targeted genes) were analysed in paraffin embedded breast cancer tissue biopsies taken from the central and peripheral parts of the tumour and were compared between them and with non-neoplastic breast tissue. Finally, **4)** comparison of PCGT methylation levels between morphologically normal tissue adjacent to the tumour and non-neoplastic breast tissue was performed.

5.2 Materials and Methods

5.2.1 Subjects

The samples were collected at the Department of Pathology, Paracelsus Private Medical University Salzburg (Salzburg, Austria). Clinical and pathologic data were stored in a database in accordance with hospital privacy regulations.

Tissue samples were collected from 50 postmenopausal women undergoing surgery for ER positive breast cancer or benign breast changes (the final number of the samples used collected from these 50 women reduced after the quality control performed of the extracted DNA). Core biopsies were dissected from the centre of IDC and from the peripheral cancer stromal border. Relevant tissue areas from tumours at least 1cm in diameter were selected on Haematoxylin and Eosin (H&E) slides and used to guide dissection from the paraffin block. For DNA extraction, 3 mm diameter core punches were used. The samples were *a priori* separated into two sets: training (15 tissues taken from the centre of the tumour and 14 from the periphery) and validation (19 tissues taken from the centre of the tumour and 20 from the periphery). Both the cores and the slides with the tissue of interest were prepared by our collaborators in Salzburg.

A second set of samples for mRNA analysis was obtained from our collaborators. Frozen breast tissue samples were collected from 63 patients with breast cancer. The breast cancer specimens were obtained immediately after resection of the breast or lumpectomy brought to the pathologist and a part of the tissue was pulverized under cooling with liquid nitrogen and stored at -70°C.

5.2.2 Cell-lines, culture conditions and reagents

Human breast cancer cell lines BT-20, ZR-75-1, MCF7, MDA-MB-231, T-47D, and SK-BR-3 were obtained from the American Type Culture Collection (ATCC) and cultured according to their recommendations. The following cell lines were generously provided by: HBL-100 from NE Hynes, F Miescher Institute, Basle, Switzerland and Hs 578T from GC Buehring, School of Public Health, Berkley, CA, USA and were cultured in Dulbecco's Modified Eagle Medium (DMEM) (Gibco Invitrogen Corporation, Lofer, Austria) containing 10% foetal bovine serum (Biochrom AG, Berlin, Germany).

5.2.3 Gene selection

Recently our group has provided evidence for a new model of carcinogenesis. The predisposition of Polycomb Repressor Complex 2 (PRC2) targets to cancer-specific DNA hypermethylation suggests a ‘crosstalk ‘between PRC2 and de novo DNA methyltransferases in precursor cancer cells with a PRC2 distribution similar to that of ES cells. This ‘crosstalk’ may be initiated and/or facilitated by various environmental exposures, transgenerational inheritance, endocrine exposure, inflammation and by age. If a cell loses the potential to terminally differentiate as a consequence of irreversible CpG methylation, it will undergo prolonged exposure to environmental onslaught, and so, more likely to acquire those mutations and/or deletions necessary for carcinogenesis. It has also been reported that PcG targets are up to 12-fold more likely to demonstrate cancer specific promoter methylation than non-targets further supporting the hypothesis that cross talk between PcG proteins and DNMTs have the potential to lay ground for the development of cancer [280]. Based on that observation, polycomb group targeted genes (PCGT) were chosen to be analysed. The genes therefore,

were occupied at least by two of the three proteins forming the PRC2 complex; EZH2, EED and SUZ12. Some of the selected genes also belong to the HOX Family Genes (n=12) known to be functionally associated with breast cancer biogenesis [389] and some of them to be methylated in breast cancer [390]. Finally, some of the genes were also oestrogen receptor target genes [391] (n=6) (*PITX2*, *ESR1*, *PGR*, *CDH13*, *DCC*, *FLJ39739*). Decreased methylation of ER-targeted genes has been shown to be associated with breast cancer risk [307, 391] hypothesising that as a function of time and dose, cumulative oestrogen exposure during lifetime leaves an epigenetic signature in the DNA, which is associated with a postmenopausal risk to develop breast cancer [116].

5.2.4 RNA isolation and reverse transcription (RT) - PCR

Total cellular RNA was extracted by the acid guanidium thiocyanate-phenol-chloroform method and cDNA was kindly prepared by Heidi Fiegl. Primers and probe for RT PCR analysis for *NEUROD1* were purchased from Applied Biosystems (Applied Biosystems Assay ID: Hs00159598_m1) designed by Heidi Fiegl. Primers and probes for the TATA box-binding protein (TBP; a component of the DNA-binding protein complex TFIID as endogenous RNA control) were used according to Bieche et al [392]. Real-time PCR was performed using an ABI Prism 7900HT Detection System (Applied Biosystems, Foster City, CA). The standard curves were generated using serially diluted solutions of standard cDNA derived from the HBL-100 breast carcinoma cell-line.

5.2.5 DNA extraction from paraffin embedded tissue samples

QIAGEN/QIA-amp Tissue kit was used. The experimental procedure was based on the protocol provided with the kit named as ‘‘Purification of Total DNA from

Animal Tissue''. For each DNA extraction approximately 25 mg of tissue were used. The samples were transferred into Sarstedt tube where 900 µl xylene was added, vortexed vigorously, centrifuged at 13000 rpm for 5 minutes with the supernatant being removed at the end of the step. Two rounds of ethanol washing (800 µl each time) followed. The supernatant was removed and the samples were incubated at 37°C for 15 minutes until all ethanol was evaporated. The tissue pellet was re-suspended in 180 µl buffer ATL and 40 µl of proteinase K were added. The samples were vortexed and incubated at 55°C (water bath) overnight.

The following day (s), 40 µl proteinase K were added till the tissue cores were fully digested. 200 µl of Buffer AL were added and they were mixed immediately by vortexing and then incubated at 70°C for 10 minutes. 200 µl of ethanol were added and mixed thoroughly by vortexing and then centrifuged. The mixture was transferred to the DNeasy Mini Spin Column (DNMSC) placed in a 2 ml collection tube and centrifuged at 8000 rpm for 1 minute. The flow–through was pipetted again into the spin column and centrifuged. The DNMSC was placed in a new 2 ml collection tube and two washing steps with 500 µl Buffer AW1 and AW1 were carried out. The DNMSC was placed in a clean 1.5 ml microcentrifuge tube (eppendorf) where 100 µl Buffer AE were added and incubated for 5 minutes, and then centrifuged for 1 minute at 8000 rpm to elute. DNA was transferred tube and for long storage they were placed at -80°C.

5.2.6 DNA quantification-NanoDrop

NanoDrop was used for the quantification of DNA. 1 µl of the buffer that the DNA was eluted in (AE buffer for the paraffin embedded tissue samples) was pipetted

onto the NanoDrop to blank the device. To check the concentration of the samples 1 µl of undiluted DNA was used.

5.2.7 DNA Quality Test

To check the DNA quality of the paraffin embedded tissue samples, real time PCR reaction was carried out using three housekeeping genes collagen 2A1 (*COLA2A1*), actin (*ACTB*) and Glyceraldehyde 3-phosphate dehydrogenase (*GAPDH*). For each reaction 4 ng/µl were used. All the primers used for the purposes of the study were provided from (Metabion, International AG, Germany). Details regarding the primers are provided in Appendix IX.

5.2.8 M.SssI Modification

M.SssI is a CpG methylase that methylates cytosines in the context of the CpG dinucleotide using SAM as a methyl donor. M.SssI-treated DNA is used as a universally methylated reference sample in all MethyLight reactions. Repeated rounds of M.SssI treatment are beneficial for methylating the genomic DNA sample. After each round of M.SssI treatment, the purified DNA sample was bisulphite-converted and tested with a methylation-specific MethyLight reaction to determine if the methylation reaches a plateau. This M.SssI-DNA sample was also used as a template for the standard curve samples.

Components

621 µl DNA (0.05 µg/µl final concentration), 100 µl 32 mM SAM (0.16 mM final concentration), 200 µl 10 X NEB2 Buffer (1 X final concentration), 50 µl M.SssI enzyme (4 units/µl) l (0.05 units/µl final concentration), water: to 2000 µl. The reaction components were mixed thoroughly and incubated at 37°C overnight.

Next day, to a volume of 2000 µl the following components were added: 28 µl SAM (1:10; 3.2 mM), 16 µl M.SssI enzyme (4units/µl) and incubated at 37°C overnight.

5.2.9 Bisulphite Modification

BM is a 3 step process involving:

1) Sulphonation 2) Hydrolytic Deamination 3) Desulphonation

BM and recovery of bisulphite-converted DNA steps were performed using the Zymo EZ DNA methylation kit (Zymo Research, Orange, CA) according to the manufacturer's instructions. M.SssI-modified DNA was also treated for use as a methylated reference in MethyLight assays. 40 µl of sample DNA were added manually to the Conversion Plate and 130 µl CT conversion reagent was added and mixed by pipetting. A foil cover was placed onto the plate and the samples were incubated overnight in a thermal cycler under the following conditions:
Program: 98°C for 10 minutes, 53°C for 30 minutes, 8 cycles: 53 °C for 6 minutes, 37°C for 30 minutes, 4°C storage for up to 6 hours

Next day, the Zymo Spin I-96 Deep well Filter Plate was placed onto a Collection Plate. M-Binding Buffer (600 µl) was added to the Zymo Spin I-96 Deep well Filter Plate. The samples were transferred to the Conversion Plate manually and the plate was centrifuged at 3000 rpm for 5 minutes. The flow through by the end of this step was discarded. The M-Desulphonation Buffer (200 µl) was added and the plate was covered and left to incubate for 15 minutes. After incubation a centrifugation step was followed at 3000 g for 5 minutes and eventually the flow through was discarded.

Two steps of washing followed by adding M-Wash Buffer (400 µl). For the first washing a centrifugation at 3000 rpm for 5 minutes was carried out and for the second washing centrifugation was at 3000 rpm for 10 minutes. Finally the collection plate was discarded and the Zymo Spin I-96 Deep well Filter Plate was placed onto an Elution Plate where 30µl of M-Elution Buffer were added. The samples were incubated for 5 minutes at room temperature and then centrifuged at 3000 rpm for 3 minutes to elute the DNA. The plate was stored at 4°C before recording eluted volumes and transferring to labelled tubes. The final tubes were stored at -80°C until required for MethyLight. The samples are prepared using the appropriate concentration by diluting them. Negative controls (water) were prepared having the same volume as the samples to be able to check that the modification was free of contamination. The positive control was Sss1-treated human peripheral blood lymphocyte DNA in order to check if the bisulphite modification was successful.

Preparation of solutions

Proteinase K: 260 µl of Proteinase K Storage Buffer had to be added to the tube containing Proteinase K, which was dissolved and then stored at -20°C.

CT Conversion Reagent: The CT Conversion Reagent was prepared prior to first use. 790 µl of M-Solubilisation Buffer and 300 µl of M-Dilution Buffer were added to a tube of CT Conversion Reagent that was mixed at room temperature with frequent vortexing for 10 minutes. Then 160 µl of M-Reaction Buffer were added and mixed for an additional 1 minute.

M-Wash Buffer: 100% ethanol of 24 ml was added to the 6 ml M-Wash Buffer before use.

5.2.10 MethyLight Primers and Probes

All the primers and probes were ordered from (Metabion, International AG, Germany). Primers and probes specific for methylated DNA and used for MethyLight reactions are listed in the Appendix IX. All primers and probes were designed by Heidi Fiegl and Sophia Apostolidou kindly showed me how to design them.

Design

The following sites were used in order to design the primers and probes for the study:

1) Sequence identification: (http://www.gene.ucl.ac.uk/nomenclature/)
2) Search of CpG islands in the promoter region: CpGplot software (http://www.ebi.ac.uk/emboss/cpgplot/), CpG Island Searcher software (http://www.uscnorris.com/cpgisland2/cpg.aspx)
3) Primer design: ABI Primer expresses software

Requirements

1) Primers were designed as close as possible to the probe without overlapping the probe (Primers)
2) The G-C content was kept in the 30 - 80% range (Primers and Probes)
3) Runs of an identical nucleotide were avoided. This is especially true for guanine (Gs), where runs of four or more Gs should be avoided (Primers and Probes)
4) When using Primer Express software, the Tm was 58 - 60 °C (Primers)
5) The five nucleotides at the 3’ end had no more than two G and/or C bases (Primers)

6) Gs on the 5’ end were avoided (Probes)
7) The strand that gives the probe more Cs than Gs was selected (Probes)
8) Temperature was at 68-70 °C using the Primer Express Software

Preparation

All primer/probe sets used were diluted to the same stock concentrations to standardize the PCR reaction set-up as well as the running of the PCR program. The primers and probes, since they are lyophilized after synthesis, need to be dissolved in sterile water. The forward and reverse primers were prepared at a concentration of 300 µM and the probe at a concentration of 100 µM.

Small aliquots of the primers at these concentrations were made to prevent repeated freeze/thaw events. The primers/probe were diluted to a working stock of 6 µM (primers) and 2 µM (probe). This is achieved by combining the stock solutions of the forward primer, reverse primer and probe in one tube as an Oligo Mix: (4 µl of the 300 µM forward primer, 4 µl of the 300 µM reverse primer, and 4 µl of the 100 µM probe in a 600 µl total volume). For a 15µl MethyLight reaction we used 2.25 µl of the Oligo Mix. This 2.25 µl volume represents the combined volumes from each of the two individual 6 µM primers and the 2 µM probe. It should be noted that the probe for each methylation reaction contains a black hole quencher (BHQ-1) at the 3’ end and a 6 FAM fluorophore at the 5’ terminus.

5.2.11 MethyLight

The technical core of DNA methylation analysis for our group is MethyLight. A sensitive, fluorescence-based real-time PCR technique that is capable of quantitating DNA methylation at a particular locus by using DNA oligonucleotides

that anneal differentially to bisulphite converted DNA according to the methylation status in the original genomic DNA. Compared with any other currently available technique for CpG methylation analysis (e.g. MSP, pyrosequencing, microarrays, etc.), MethyLight offers the unique opportunity to be run as a highly sensitive high throughput facility while being able to provide quantitative measures without the necessity for manipulation of post-PCR products. A single disadvantage of MethyLight is that it cannot detect allele-specific methylation differences (which would be possible by BM pyrosequencing), which is not needed for the purpose of this project.

The specificity of the reactions for methylated DNA was confirmed separately by using in all MethyLight reactions SssI-treated human peripheral blood lymphocyte DNA (New England Biolabs, UK) which is prepared in advance and used as a universally methylated reference sample. One set of primers and probes for *COL2A1* was designed specifically for bisulphite-converted DNA and used as reference set to normalize for input DNA. The percentage of fully methylated molecules at a specific locus was calculated by dividing the *GENE: COL2A1* ratio of a sample by the *GENE*: *COL2A1* ratio of *Sss*I-treated WBC DNA and multiplying by 100. The abbreviation PMR indicates this measurement. A gene was deemed methylated if the percentage of fully methylated reference value was >0.

Conduction of MethyLight assay

The MethyLight assay utilises the TaqMan PCR principle which requires forward and reverse primers as well as an oligomeric probe which emits fluorescence only after it is degraded by the 5'-3' exonuclease activity of Taq polymerase

without uracil DNA glycosylase (Applied Biosystems). Each PCR reaction uses the same basic reaction set-up. Uracil DNA glycosylase (AMPerase) was not included as a component in the PCR reactions. AMPerase catalyzes the removal of uracil, and this is problematic since bisulphite converted DNA is used as a DNA template and will therefore contain uracil (from unmethylated cytosines). After the primer/probe preparation the MasterMix Reactions were prepared mixing the OligoMix (2.25 µl) and TaqMan PCR (7.5 µl) at a final volume of 10µl (adding water). Finally, 5 µl of the bisulphite DNA sample were disposed into the wells of a 384-well plate and then 10 µl of the MasterMix were added. The plate was sealed, mixed and centrifuged at 1500 rpm for 1 minute. Then it was placed in real-time PCR instrument.

Program:

95°C for 10 minutes, then 50 cycles of: 95°C for 15 seconds, 60°C for 1 minute

Bisulphite converted M.SssI-treated DNA was also included in different concentrations used as a standard curve for each methylation and control reactions (water and OligoMix).

5.2.12 Statistics

Statistical analysis was carried out using a computer assisted program-SPSS version 12.0.1, Chicago, IL. For both training and validation sets for each gene the percent of non-zero results, the median and the p-values from the Mann-Whitney test was calculated. The genes in both training and validation sets were also assessed using Receiver Operator Characteristic (ROC) curves and the Area under the Curve (AUC) value. When a gene was denoted a ‘reverse

decision rule' was applied if higher methylation values meant it was more likely the subject to be a control rather than cancer case. In order to assess if there is a difference between the centre and periphery groups, a non-parametric paired test (Wilcoxon rank test) was carried out comparing the rank order of values for centre versus periphery. Spearman correlation analysis was performed in order to examine any association between the two zones.

5.3 Results

5.3.1 Clinicopathological characteristics of the study subjects

The samples taken from the study women were divided into two sets: training and validation. The training set consisted of 25 cases and 25 controls and the validation set of 25 cases and 25 controls. The number of cases and controls in both sets dropped after the DNA quality test. As it is seen in Table 5-1 there were no statistically significant differences (p-value less than 0.05) between the clinicopathological features of the two sets.

Table 5-1: Clinicopathological features of the study women for training and validation set for the analysis of (A) the tissue taken from the centre and periphery and (B) the morphologically normal tissue adjacent to the tumour.

(A)

Clinicopathological Features		Training Set (15)	Validation Set (20)	p-value
Mean Age		60	59.55	0.298
Histological Type	**IDC** **IDC+DCIS**	5 10	8 12	0.227
Grading	**I** **II** **III**	1 10 4	0 17 3	0.073
Staging	**1** **2**	11 4	11 9	0.096
PR	**positive** **negative**	12 3	15 5	0.237
HER2 positive	**1** **2** **0**	7 1 7	5 3 12	0.121
Sentinel Node	**positive** **negative**	5 15	9 11	0.072

DCIS=ductal carcinoma *in situ*; IDC=invasive ductal carcinoma; HER2=human epidermal growth factor receptor 2; PR=progesterone receptor

(B)

Clinicopathological Features		Training Set (19)	Validation Set (20)	p-value
Mean Age		60	60.75	0.287
Histological Type	**IDC** **IDC+DCIS**	5 14	9 11	0.107
Grading	**I** **II** **III**	1 13 5	0 17 3	0.132
Staging	**1** **2**	13 6	11 9	0.243
PR	**positive** **negative**	15 4	16 4	0.803
HER2+	**1** **2** **0**	8 3 8	5 2 13	0.121
Sentinel Node	**positive** **negative**	6 13	6 14	0.789

DCIS=ductal carcinoma *in situ*; IDC=invasive ductal carcinoma; HER2=human epidermal growth factor receptor 2; PR=progesterone receptor

The control samples were taken from postmenopausal women that were undergoing surgery for benign conditions and were age matched to the women diagnosed with breast cancer.

5.3.2 DNA quality test

Specific criteria were set up to select the eligible samples to perform the analysis. The quality of the genomic DNA was checked by two methods 1) quantification (Nano-Drop) and 2) real-time PCR using three housekeeping genes collagen (*COL2A1)*, actin (*ACTB)* and glyceraldehyde-3-phosphate dehydrogenase (*GAPDH)*. The samples that had sufficient DNA for the purposes of the analysis and a mid-exponential cycle threshold (Ct) value of <36, were included in the study. The samples that did not meet the above criteria were excluded.

5.3.3 *NEUROD1* methylation and mRNA expression

Based on two of our recent studies, we showed that *NEUROD1* was one of the best discriminators between breast cancer and non neoplastic tissue samples [280] and that methylated *NEUROD1* promoter could be a good predictive marker for chemosensitivity in breast cancer [305]. Since cancer specific methylation of some of the PCGT genes, such as *MYOD1* and *NEUROG1* has been shown not to be associated with expression in epithelial cancers [280] we wanted to investigate the association of the methylation and expression profile of *NEUROD1* in breast cancer. Therefore, we analysed and compared *NEUROD1* methylation (a specific sequence within CpG island in the promoter region of the gene was chosen based on our previous publication to be analysed, details are provided in Appendix Table IX.1 and 2) (Figure 5-2) and *NEUROD1* mRNA expression (its expression was controlled for collagen) (Figure 5-3) in a panel of 9 human breast cancer cell lines. As it is illustrated in Figure 5-3, from the 9 tested cell lines only one expressed *NEUROD1* whereas *NEUROD1* methylation was observed in 8 out of 9 cell lines. In the only cell line that expression was observed was HBL-100 which is established from human breast milk and possibly not the most representative breast cancer cell line. To further examine this observation, we analysed *NEUROD1* expression and methylation of 63 frozen breast cancer tissue samples. The experiments were performed in triplicate and, when at least two signals of the analysis had given a negative value, the expression was considered as zero. As is seen in Table 5-2, the majority of the cases; 54/63 (85.7%) did not express *NEUROD1* in contrast to *NEUROD1* methylation, which was detected in all cases with a PMR value ranging from 0.047-632.995.

Figure 5-2: *NEUROD1* DNA methylation of the different cell lines.

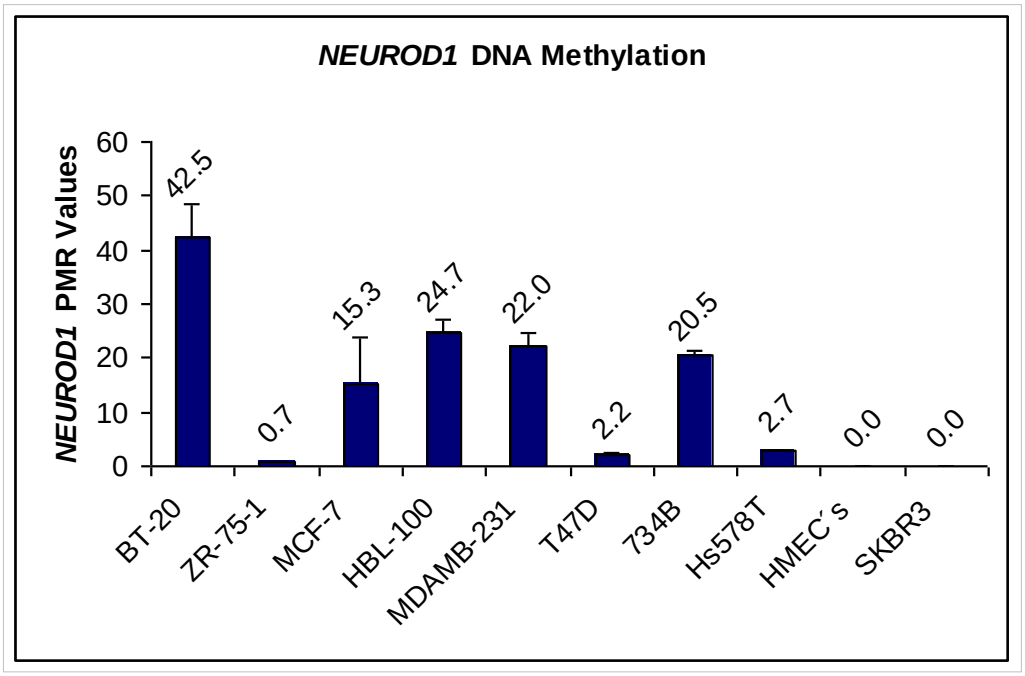


Figure 5-3: *NEUROD1* mRNA expression of the different cell lines.

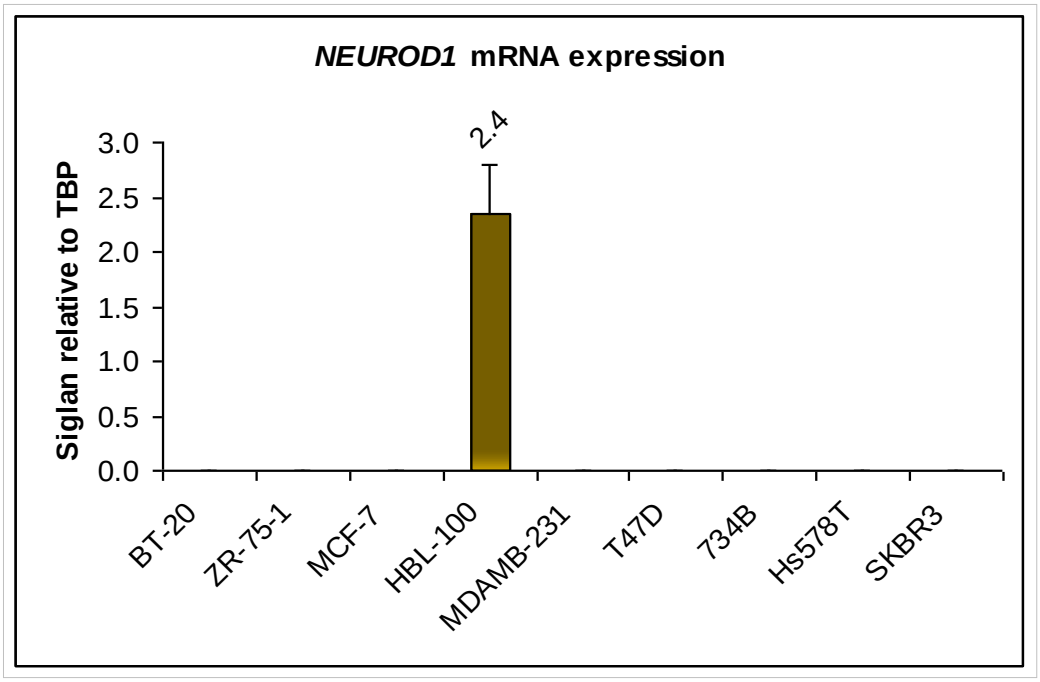


Table 5-2: *NEUROD1* methylation and mRNA expression in 63 breast cancer specimens.

Case	*NEUROD1*		Case	*NEUROD1*	
	PMR	Expression		PMR	Expression
1	0.047	0.000	33	25.456	0.000
2	0.246	0.000	34	29.153	0.000
3	0.641	0.000	35	29.829	0.000
4	0.669	2.595	36	30.277	0.000
5	0.708	0.000	37	32.528	0.000
6	0.722	0.000	38	34.396	0.000
7	0.892	93.856	39	34.549	0.302
8	1.495	0.000	40	36.433	0.000
9	1.912	0.000	41	36.685	0.000
10	3.205	0.000	42	39.516	0.000
11	3.628	0.000	43	46.768	0.000
12	4.412	0.000	44	46.908	0.000
13	4.898	0.000	45	47.793	0.000
14	5.067	0.000	46	52.011	0.000
15	6.855	0.000	47	56.516	1.064
16	6.875	0.000	48	59.441	0.416
17	6.889	0.000	49	61.509	0.000
18	7.02	0.000	50	63.278	0.000
19	7.234	0.000	51	71.241	2.882
20	9.178	0.000	52	75.363	0.000
21	9.866	0.000	53	75.382	0.000
22	10.85	0.000	54	75.608	1.959
23	12.713	0.000	55	79.196	0.000
24	12.82	0.000	56	79.849	0.000
25	13.714	0.000	57	84.717	0.000
26	15.965	0.000	58	85.034	0.000
27	16.285	0.000	59	86.433	0.000
28	17.025	0.849	60	86.963	0.000
29	18.225	0.000	61	95.934	0.000
30	19.358	0.000	62	202.137	0.000
31	23.485	1.125	63	482.26	0.000
32	23.731	0.000			

PMR= percentage of methylated reference

5.3.4 DNA methylation of PCGT genes and breast cancer

In order to further investigate the methylation level changes of the PCGT genes in breast cancer the sample sets consisting of tumour taken from the centre and from the periphery from the 35 postmenopausal women with ER positive breast cancer were compared to these with non neoplastic breast tissue taken from 40 postmenopausal women who had under gone surgery for benign breast changes. These are not the ideal control samples and another possibility would have been to have tissue from healthy postmenopausal women undergoing mammoplasty reduction–still with this approach though problems arise in getting enough progenitor cells for the purposes of the analysis. In the training set, methylation of 55 PCGT genes in 14 tumour tissues taken from the centre and 15 tumour tissues taken from the periphery and 22 controls were analysed. 24% (13 out of 55) of the genes were cancer specific ($p<0.05$), being more frequently methylated in tumour samples compared to non-neoplastic tissues as illustrated in Table 5-3.

To test the hypothesis that the selected genes are cancer predictors, they were assessed by ROC analysis as seen in Table 5-4. In this table for both the tumour taken from the centre and periphery groups the p-values given reflect whether the AUCs are significantly different from 0.5 (a straight line from bottom left to top right corners, implying a decision rule no better than chance). The predictive value of the 13 genes tested by ROC analysis showed a range of 0.71-0.95. To further validate these findings an independent validation set consisting of 19 tumour tissues taken from the centre, 20 tumour samples taken from the periphery and 18 controls were analysed examining the 13 PCGT genes from the training set that had a $p<0.05$ in the Mann-Whitney analysis for both samples taken from the centre and the periphery of the tumour. All 13 genes (Table 5-5)

were confirmed and were shown to be statistically significant ($p<0.05$). We calculated the predictive potential of these 13 genes by ROC analysis as illustrated in Table 5-6. Interestingly, even though the sample size is low, the results were consistent and the same panel of genes that were shown to be statistically significant with the Mann-Whitney test also had a statistically significant AUC value ($p<0.05$).

Table 5-3: Summary statistics of controls versus tumour tissue taken from the centre and periphery from the breast cancer cases analysed in the training set. P-value from the Mann-Whitney test for each gene is provided (significant p-value less than 0.05). The genes were ordered according to the rank of the p-value for the test of TUC versus control, and a further column for the TUP group gave the rank order value for the TUP versus control group to facilitate comparison

Genes name	Training Set								
	Control (n = 22)		Centre (n = 15)			Periphery (n = 14)			
	% positive	Median PMR *	% positive	Median PMR	p value	% positive	Median PMR	p value	p-value Rank
HOXD9	77.30%	0.01	100.00%	1.42	**0.000**	100.00%	2.48	**0.000**	1
HOXA7	90.90%	0.28	100.00%	5.39	**0.000**	100.00%	3.96	**0.000**	3
PENK	9.10%	0.00	73.30%	0.09	**0.000**	78.60%	4.22	**0.000**	2
TMEFF2	50.00%	0.00	93.30%	8.96	**0.000**	92.90%	7.06	**0.000**	5
HOXA1	59.10%	0.03	80.00%	11.74	**0.002**	85.70%	11.31	**0.000**	6
MT1A	95.50%	69.92	100.00%	28.72	**0.003**	100.00%	19.20	**0.003**	15
CRABP1	0.00%	-	33.30%	0.00	**0.004**	42.90%	0.00	**0.001**	9
GATA4	22.70%	0.00	60.00%	0.11	**0.007**	28.60%	0.00	0.609	35
HOXD11	63.60%	0.75	80.00%	17.53	**0.007**	92.90%	35.48	**0.000**	8
HOXD12	54.50%	0.03	73.30%	10.12	**0.010**	85.70%	7.75	**0.002**	12
NEUROD1	54.50%	0.00	73.30%	0.39	**0.020**	78.60%	5.19	**0.002**	13
GAD1	100.00%	0.72	100.00%	2.57	**0.020**	100.00%	4.33	**0.000**	7
HOXA13	54.50%	1.14	66.70%	177.49	**0.021**	100.00%	181.40	**0.000**	4
PITX2 (II)	72.70%	0.26	73.30%	4.33	**0.029**	85.70%	10.64	**0.001**	10

PMR= percentage of methylated reference

Genes name	Training Set								
	Control (n = 22)		Centre (n = 15)			Periphery (n = 14)			
	% positive	Median PMR *	% positive	Median PMR	p value	% positive	Median PMR	p value	p-value Rank
HIC1	63.60%	28.33	100.00%	42.58	0.043	100.00%	31.32	0.054	22
PGR	23.80%	0.00	0.00%	-	0.045	21.40%	0.00	0.680	39
HOXD8	63.60%	0.18	80.00%	4.57	0.052	64.30%	7.04	0.111	28
ITGA4	0.00%	-	13.30%	0.00	0.083	28.60%	0.00	0.009	17
PITX2 (I)	0.00%	-	13.30%	0.00	0.083	28.60%	0.00	0.009	18
CACNA1G	0.00%	-	13.30%	0.00	0.083	14.30%	0.00	0.072	24
TWIST	13.60%	0.00	33.30%	0.00	0.086	14.30%	0.00	0.829	44
MT3	77.30%	0.07	46.70%	0.00	0.107	57.10%	0.00	0.041	21
EBF3(DKFZ)	4.50%	0.00	20.00%	0.00	0.121	42.90%	0.00	0.004	16
GABRA2	4.50%	0.00	20.00%	0.00	0.136	21.40%	0.00	0.115	29
DLC1	13.60%	0.00	0.00%	-	0.142	14.30%	0.00	0.957	50
GATA5	13.60%	0.00	33.30%	0.00	0.144	50.00%	0.02	0.010	19
HOXA11	100.00%	26.31	100.00%	19.15	0.146	100.00%	19.77	0.092	27
CDH13	13.60%	0.00	33.30%	0.00	0.156	14.30%	0.00	1.000	55
SFRP4	50.00%	0.00	33.30%	0.00	0.171	21.40%	0.00	0.057	23
NEUROD2	59.10%	0.02	60.00%	2.34	0.189	85.70%	7.14	0.001	11
ESR1	50.00%	0.00	20.00%	0.00	0.201	14.30%	0.00	0.024	20
HOXB7	77.30%	0.02	53.30%	0.00	0.202	78.60%	0.02	0.636	37
NEUROG1	0.00%	-	6.70%	0.00	0.226	14.30%	0.00	0.072	26
HOXA6	59.10%	2.21	40.00%	0.00	0.230	42.90%	0.00	0.410	33

PMR= percentage of methylated reference

Genes name	Training Set								
	Control (n = 22)		Centre (n = 15)			Periphery (n = 14)			
	% positive	Median PMR *	% positive	Median PMR	p value	% positive	Median PMR	p value	p-value Rank
SLC6A20	9.10%	0.00	0.00%	-	0.236	7.10%	0.00	0.892	47
SFRP1	45.50%	0.00	26.70%	0.00	0.242	35.70%	0.00	0.885	46
ZBTB16	54.50%	0.00	33.30%	0.00	0.266	28.60%	0.00	0.269	31
HOXA10	95.50%	2.91	80.00%	7.13	0.300	85.70%	5.03	0.626	36
DCC	4.50%	0.00	13.30%	0.00	0.311	7.10%	0.00	0.713	41
SLIT2	63.60%	0.12	46.70%	0.00	0.317	50.00%	0.02	0.637	38
IGF2	4.50%	0.00	13.30%	0.00	0.343	7.10%	0.00	0.713	40
HOXC9	18.20%	0.00	6.70%	0.00	0.376	7.10%	0.00	0.331	32
GDNF	18.20%	0.00	6.70%	0.00	0.376	14.30%	0.00	0.764	43
TITF1	18.20%	0.00	6.70%	0.00	0.376	14.30%	0.00	0.920	48
HOXA9	100.00%	11.88	86.70%	7.74	0.404	100.00%	11.52	1.000	54
CYP27B1	100.00%	4.86	100.00%	4.65	0.458	0.00%	6.34	0.183	30
MYOD1	22.70%	0.00	26.70%	0.00	0.498	64.30%	0.21	0.003	14
SFRP5	81.80%	2.43	60.00%	0.63	0.708	71.40%	1.54	0.744	42
CYP1B1	4.50%	0.00	6.70%	0.00	0.752	0.00%	-	0.425	34
CALCA	45.50%	0.00	40.00%	0.00	0.784	35.70%	0.00	0.942	49
FLJ39739	50.00%	0.00	33.30%	0.00	0.973	35.70%	0.00	0.873	45
GATA3	0.00%	-	0.00%	-	1.000	14.30%	0.00	0.072	25
PYCARD	0.00%	-	0.00%	-	1.000	0.00%	-	1.000	51
TP73	0.00%	-	0.00%	-	1.000	0.00%	-	1.000	52
BCL22	0.00%	-	0.00%	-	1.000	0.00%	-	1.000	53

PMR= percentage of methylated reference

Table 5-4: ROC analysis for both samples taken from the centre and periphery of the tumour in training set. The performance of each gene as a predictor of breast cancer was assessed using ROC curves and the AUC value. The genes were ordered according to the rank of the p-value for the test of the tissue samples taken from the center and periphery of tumour versus control. Significance required a p-value of less than 0.05. * 'Reverse decision rule' applied means that the higher the methylation value, the more likely the subject is a control rather than breast cancer case (tumour taken from the center and periphery).

Genes name	Training Set								
	Centre				Periphery				
	AUC	95% CI Lower	95% CI Upper	p value	AUC	95% CI Lower	95% CI Upper	p value	p-value Rank
HOXD9	0.930	0.852	1.008	**0.000**	0.955	0.894	1.015	**0.000**	5
HOXA7	0.891	0.788	0.993	**0.000**	0.935	0.859	1.011	**0.000**	11
TMEFF2	0.853	0.719	0.987	**0.000**	0.888	0.767	1.009	**0.000**	9
PENK	0.830	0.682	0.979	**0.001**	0.883	0.745	1.021	**0.000**	10
HOXA1	0.792	0.619	0.966	**0.003**	0.867	0.715	1.019	**0.000**	1
**MT1A*	0.788	0.051	0.373	**0.003**	0.795	0.048	0.361	**0.003**	13
HOXD11	0.761	0.584	0.937	**0.008**	0.860	0.726	0.994	**0.000**	6
HOXD12	0.745	0.564	0.927	**0.012**	0.808	0.650	0.967	**0.002**	12
GAD1	0.727	0.557	0.897	**0.020**	0.870	0.733	1.007	**0.000**	2
NEUROD1	0.721	0.540	0.902	**0.024**	0.805	0.633	0.977	**0.002**	4
HOXA13	0.718	0.526	0.910	**0.026**	0.922	0.835	1.009	**0.000**	3
PITX2 (II)	0.712	0.513	0.911	**0.030**	0.825	0.656	0.993	**0.001**	7
CRABP1	0.667	0.477	0.856	**0.089**	0.714	0.526	0.903	**0.032**	8

AUC= Area under the curve; CI= confidence interval

Table 5-5: Percentage of positive cases and distribution of methylation levels of the 13 genes tested in validation set. Controls and breast cancer cases (tumour samples taken from the centre and periphery) showing the percentage of positive cases and the median PMR values. P-values are provided from Mann-Whitney test for each gene. Significance required a p-value of less than 0.05 after Mann-Whitney test.

Genes name	Validation Set								
	Controls (n = 18)		Centre (n = 19)			Periphery (n = 20)			
	% positive	Median PMR	% positive	Median PMR	p value	% positive	Median PMR	p value	p-value Rank
HOXA1	44.40%	0	94.70%	26.15	**0.000**	89.50%	25.54	**0.000**	3
GAD1	100.00%	0.78	100.00%	15.00	**0.000**	94.70%	14.72	**0.000**	2
HOXA13	38.90%	0	94.70%	152.61	**0.000**	84.20%	126.08	**0.000**	5
CRABP1	0.00%	-	73.70%	5.16	**0.000**	57.90%	3.12	**0.000**	7
NEUROD1	44.40%	0	89.50%	9.73	**0.000**	89.50%	6.51	**0.000**	4
HOXD9	100.00%	0.24	100.00%	8.02	**0.000**	100.00%	1.84	**0.002**	10
PITX2(II)	38.90%	0	84.20%	5.69	**0.000**	84.20%	4.25	**0.000**	6
HOXD11	66.70%	0.1	94.70%	39.07	**0.000**	84.20%	32.94	**0.001**	9
TMEFF2	83.30%	0.15	89.50%	22.34	**0.000**	100.00%	16.57	**0.000**	1
PENK	22.20%	0	68.40%	0.29	**0.001**	73.70%	0.85	**0.000**	8
HOXA7	100.00%	1.86	94.70%	6.22	**0.004**	100.00%	4.31	**0.025**	12
HOXD12	66.70%	0.17	78.90%	4.98	**0.006**	89.00%	6.97	**0.002**	11
MT1A	100.00%	73.06	100.00%	37.55	**0.023**	100.00%	46.07	**0.027**	13

PMR= Percentage of Methylated Reference

Table 5‑6: ROC analysis for both tumour samples taken from centre and periphery in validation set. The performance of each gene as a predictor of cancer was assessed using ROC curves and the AUC value. The significant assessment required a p-value of less than 0.05. The genes were ordered according to the rank of the p-value for the test of samples taken from the centre and periphery of the tumour versus control. * 'Reverse decision rule' applied means that the higher the methylation value, the more likely the subject is a control rather than cancer case.

Genes name	Validation Set								
	Centre				Periphery				
	AUC	95% CI Lower	95% CI Upper	p value	AUC	95% CI Lower	95% CI Upper	p value	p-value Rank
HOXD9	0.889	0.787	0.991	**0.000**	0.801	0.659	0.944	**0.002**	9
HOXA7	0.778	0.626	0.93	**0.004**	0.716	0.551	0.881	**0.025**	12
TMEFF2	0.86	0.722	0.997	**0.000**	0.977	0.936	1.017	**0.000**	1
PENK	0.792	0.642	0.943	**0.002**	0.822	0.68	0.963	**0.001**	8
HOXA1	0.950	0.872	1.029	**0.000**	0.901	0.788	1.014	**0.000**	3
**MT1A*	0.719	0.111	0.450	**0.020**	0.713	0.637	0.942	**0.027**	13
HOXD11	0.883	0.767	0.999	**0.000**	0.825	0.679	0.971	**0.001**	7
HOXD12	0.76	0.598	0.923	**0.007**	0.789	0.638	0.941	**0.003**	11
GAD1	0.936	0.853	1.018	**0.000**	0.918	0.810	1.026	**0.000**	2
NEUROD1	0.904	0.795	1.012	**0.000**	0.892	0.78	1.004	**0.000**	4
HOXA13	0.925	0.837	1.014	**0.000**	0.858	0.729	0.987	**0.000**	5
PITX2 (II)	0.876	0.752	0.999	**0.000**	0.852	0.723	0.981	**0.000**	6
CRABP1	0.868	0.742	0.994	**0.000**	0.789	0.637	0.942	**0.003**	10

AUC= Area under the curve; CI= confidence interval

5.3.5 DNA methylation of PCGT genes and intra-tumour heterogeneity

Even though we show methylation levels of specific genes to be comparable between selected tissue samples that were taken from the center and the periphery of the tumour in both the training and validation sets, there were also some genes that demonstrated differential methylation levels between the two tumour zones. This was true for both cancer specific genes (genes that were shown to be specifically methylated in the breast tumour tissue when they were compared to the controls) but also for the genes that were shown to be non-specifically methylated in cancer. In order to verify these results we performed non-parametric paired test for the genes one by one. Comparison between the two zones within the tumour for cancer specific genes (Table 5-7) and non-cancer specific genes (Table 5-8) did not show any difference in the methylation levels. In order to investigate whether there is any correlation between the two different zones we performed Spearman correlation analysis. The analysis showed 4 out of the 13 cancer specific genes to be positively correlated in both the training and validation for the two different zones of the tumour (Table 5-7).

Table 5-7: Comparison of DNA methylation changes between the samples that were taken from the centre and periphery of the tumour with non-parametric paired test and Spearman correlation analysis for the breast cancer specific genes. P-values are given from the 13 genes that were further confirmed in the validation set. Non-parametric paired test analysis comparing the rank order of values for tissue taken from the centre of the tumour versus tissue taken from the periphery of the tumour to assess if there is a difference in their methylation levels and correlation coefficient analysis is also demonstrated.

Breast Cancer Specific Genes	Non-parametric paired test p-value	Spearman Correlation coefficient p-value	
HOXD9	0.675	0.339	0.054
HOXA7	0.993	0.120	0.507
PENK	0.766	0.110	0.544
TMEFF2	0.280	0.399	0.021
HOXA1	0.614	0.347	0.048
MT1A	0.714	0.675	0.000
CRABP1	0.715	0.567	0.001
HOXD11	0.822	0.313	0.077
HOXD12	0.217	0.216	0.227
NEUROD1	0.814	0.361	0.039
GAD1	0.526	0.266	0.135
HOXA13	0.382	-0.035	0.847
PITX2 (II)	0.829	0.256	0.151

Table 5-8: Comparison of DNA methylation changes in non-cancer specific genes between tissue taken from the centre of the tumour and tissue taken from the periphery of the tumour with non-parametric paired test analysis.

Non Breast Cancer Specific Genes	p-value
SFRP1	0.063
NEUROD2	0.099
PGR	0.109
GATA5	0.114
ESR1	0.144
DLC1	0.180
GATA3	0.180
GDNF	0.180
MYOD1	0.203
CACNA1G	0.285
CDH13	0.310
CYP1B1	0.317
SLC6A20	0.317
EBF3 (DKFZ)	0.398
ZBTB16	0.401
MT3	0.424
HOXA11	0.470
HOXA6	0.484
GABRA2	0.500
SFRP5	0.534
HOXA10	0.551
CYP27B1	0.594
DCC	0.655
HOXC9	0.655
IGF2	0.655
NEUROG1	0.655
TITF1	0.655
CALCA	0.674
FLJ39739	0.674
HOXB7	0.730
SLIT2	0.790
ITGA4	0.893
PITX2 (I)	0.893
TWIST	0.893
SFRP4	0.917
HOXA9	0.925
HOXD8	0.972
BCL2	1.000
PYCARD	1.000
TP73	1.000

5.3.6 PCGT gene methylation in the normal tissue adjacent to the tumour and breast cancer risk prediction

In order to investigate DNA methylation changes of the PCGT genes in morphologically normal tissue adjacent to the breast tumour a set consisting of 19 postmenopausal women with ER+ breast cancer (cases) were analysed and compared with non neoplastic tissue from 22 postmenopausal women who had undergone surgery for benign breast changes (controls). In the training set, methylation changes of 55 PCGT genes were examined. 5.5% (3 genes out of 55; *HOXD8*, *SL6A20* and *HOXA9*) of the genes analysed showed significant ($p<0.05$) methylation changes in the normal tissue adjacent to the tumour when compared with the non neoplastic tissue as it is illustrated in Table 5-9. The majority of the genes did not show any methylation changes.

To further validate these findings, we analysed in an independent set consisting of 20 cases and 20 controls the identified genes from the training set which had a $p<0.05$. One of the three genes analysed, *HOXA9* was confirmed and was shown to be statistically significant in the validation set as illustrated in Table 5-10. Interestingly, the results for *HOXA9* were consistent with the median PMR values observed in the training set for both cases and controls being also observed in the validation set. In addition, it is worth noting that *HOXA9* demonstrated less frequent methylation in the normal tissue adjacent to the tumour compared to the controls. To test the hypothesis that *HOXA9* is a breast cancer predictor, ROC analysis for both training and validation set was performed. The predictive value of *HOXA9* was statistically significant showing an AUC value of 0.677 ($p=0.05$) for the training set and an AUC value of 0.682 ($p=0.048$) for the validation set.

Table 5-9: Summary statistics of controls versus normal tissue adjacent to the breast tumour analysed in the training set. P-value from the Mann-Whitney test for each gene is provided (significant p-value less than 0.05).

Genes name	Training Set								
	Controls				Normal Tissue Adjacent to Breast Tumour				
	% positive	Median PMR	Minimum	Maximum	% positive	Median PMR	Minimum	Maximum	p-value
HOXD8	63.6%	0.18	0.00	15.60	21.1%	0.00	0.00	7.30	**0.01**
SLC6A20	9.1%	0.00	0.00	0.04	31.6%	0.00	0.00	5.15	**0.05**
HOXA9	100.0%	11.88	0.01	28.74	100.0%	7.01	1.41	44.32	**0.05**
GDNF	18.2%	0.00	0.00	4.64	0.0%	0.00	0.00	0.00	0.06
HOXA6	59.1%	2.21	0.00	19.21	26.3%	0.00	0.00	23.65	0.06
CALCA	45.5%	0.00	0.00	2.40	15.8%	0.00	0.00	2.39	0.07
FLJ39739	50.0%	0.00	0.00	0.01	21.1%	0.00	0.00	0.03	0.11
HOXA1	59.1%	0.03	0.00	12.49	26.3%	0.00	0.00	11.67	0.12
GATA3	0.0%	0.00	0.00	0.00	10.5%	0.00	0.00	0.01	0.12
SLIT2	63.6%	0.12	0.00	4.22	47.4%	0.00	0.00	4.24	0.13
SFRP5	81.8%	2.43	0.00	16.75	57.9%	0.03	0.00	11.18	0.13
PITX2 (II)	72.7%	0.26	0.00	10.43	47.4%	0.00	0.00	3.53	0.25
CYP27B1	100.0%	4.86	1.18	10.65	100.0%	5.87	1.98	12.20	0.26
PYCARD (ASC)	0.0%	0.00	0.00	0.00	5.3%	0.00	0.00	0.00	0.28
CACNA1G	0.0%	0.00	0.00	0.00	5.3%	0.00	0.00	3.37	0.28
CRABP1	0.0%	0.00	0.00	0.00	5.3%	0.00	0.00	0.27	0.28

PMR= percentage of methylated reference

Genes name	Training Set								
	Controls				Normal Tissue Adjacent to Breast Tumour				
	% positive	Median PMR	Minimum	Maximum	% positive	Median PMR	Minimum	Maximum	p-value
TP73	0.0%	0.00	0.00	0.00	5.3%	0.00	0.00	18.55	0.28
HOXD12	54.5%	0.03	0.00	10.28	84.2%	0.99	0.00	8.43	0.31
HOXA10	95.5%	2.91	0.00	28.98	89.5%	4.25	0.00	18.92	0.35
ZBTB16	54.5%	0.00	0.00	3.30	36.8%	0.00	0.00	0.33	0.35
IGF2	4.5%	0.00	0.00	0.11	0.0%	0.00	0.00	0.00	0.35
CYP1B1	4.5%	0.00	0.00	0.11	0.0%	0.00	0.00	0.00	0.35
GATA5	13.6%	0.00	0.00	1.49	5.3%	0.00	0.00	0.04	0.36
DLC1	13.6%	0.00	0.00	0.94	5.3%	0.00	0.00	0.79	0.39
NEUROD1	54.5%	0.00	0.00	5.20	52.6%	0.33	0.00	5.15	0.39
ESR1	50.0%	0.00	0.00	4.03	31.6%	0.00	0.00	5.73	0.42
GABRA2	4.5%	0.00	0.00	0.00	10.5%	0.00	0.00	3.78	0.43
EBF3 (DKFZ)	4.5%	0.00	0.00	0.00	10.5%	0.00	0.00	18.55	0.43
DCC	4.5%	0.00	0.00	0.01	10.5%	0.00	0.00	1.43	0.43
GAD1	0.0%	0.72	0.00	28.64	94.7%	0.98	0.00	11.43	0.46
NEUROD2	59.1%	0.02	0.00	9.29	63.2%	0.17	0.00	10.14	0.53
HOXD11	63.6%	0.75	0.00	58.02	68.4%	1.26	0.00	33.40	0.54
PENK	9.1%	0.00	0.00	0.28	15.8%	0.00	0.00	1.54	0.55
HIC1	63.6%	28.33	0.00	73.74	63.2%	22.13	0.00	60.88	0.59
MT3	77.3%	0.07	0.00	3.85	73.7%	0.03	0.00	1.78	0.66

PMR= percentage of methylated reference

Genes name	Training Set								
	Controls				Normal Tissue Adjacent to Breast Tumour				
	% positive	Median PMR	Minimum	Maximum	% positive	Median PMR	Minimum	Maximum	p-value
MYOD1	22.7%	0.00	0.00	2.04	26.3%	0.00	0.00	1.48	0.68
HOXC9	18.2%	0.00	0.00	0.31	21.1%	0.00	0.00	8.28	0.68
TWIST	13.6%	0.00	0.00	0.01	15.8%	0.00	0.00	0.30	0.70
HOXB7	77.3%	0.02	0.00	1.88	63.2%	0.01	0.00	4.58	0.72
SFRP1	45.5%	0.00	0.00	16.36	42.1%	0.00	0.00	17.09	0.75
SFRP4	50.0%	0.00	0.00	58.57	47.4%	0.00	0.00	46.14	0.78
HOXA13	54.5%	1.14	0.00	242.64	47.4%	0.00	0.00	144.87	0.78
CDH13	13.6%	0.00	0.00	1.87	15.8%	0.00	0.00	0.77	0.86
HOXA7	90.9%	0.28	0.00	4.22	0.0%	0.34	0.01	5.66	0.90
TMEFF2	50.0%	0.00	0.00	6.46	47.4%	0.00	0.00	18.07	0.91
MT1A	95.5%	69.92	0.00	220.33	0.0%	67.83	19.93	129.29	0.92
HOXA11	100.0%	26.31	1.15	66.64	100.0%	27.53	4.01	71.03	0.92
GATA4	22.7%	0.00	0.00	3.23	21.1%	0.00	0.00	6.71	0.93
HOXD9	77.3%	0.01	0.00	1.78	68.4%	0.01	0.00	11.85	0.94
TITF1	18.2%	0.00	0.00	0.40	15.8%	0.00	0.00	11.67	0.97
PGR	23.8%	0.00	0.00	2.02	26.3%	0.00	0.00	5.32	0.99
BCL2	0.0%	0.00	0.00	0.00	0.0%	0.00	0.00	0.00	1.00
NEUROG1	0.0%	0.00	0.00	0.00	0.0%	0.00	0.00	0.00	1.00
ITGA4	0.0%	0.00	0.00	0.00	0.0%	0.00	0.00	0.00	1.00
PITX2 (I)	0.0%	0.00	0.00	0.00	0.0%	0.00	0.00	0.00	1.00

PMR= percentage of methylated reference

Table 5-10: Summary statistics of controls versus normal tissue adjacent to the breast tumour analysed in the validation set. P-value from the Mann-Whitney test for each gene is provided (significant p-value less than 0.05).

Genes name	Validation Set								
	Controls				Normal Tissue Adjacent to Breast Tumour				
	% positive	Median PMR	Minimum	Maximum	% positive	Median PMR	Minimum	Maximum	p-value
HOXA9	100%	**12.59**	0.34	19.03	90%	6.51	0.00	21.49	**0.05**
HOXD8	50%	0.02	0.00	14.85	35%	0.00	0.00	14.16	0.37
SLC6A20	0%	0.00	0.00	0.00	0%	0.00	0.00	0.00	1.00

PMR= percentage of methylated reference

5.3.7 Comparison of DNA methylation changes observed in breast tumour and the corresponding normal tissue

Comparison of DNA methylation changes observed between the two tissues, the morphologically normal tissue adjacent to the breast tumour and the breast tumour, was taken into account. As it was shown in Table 5-3, 13 genes showed to have significant DNA methylation changes in the breast tumour compared to controls but none of these were shown to have a significant methylation change in the corresponding morphologically normal tissue adjacent to the tumour (Table 5-9). Significant methylation changes within the corresponding normal tissue were observed only for one gene, *HOXA9* which was not shown to be one of the breast cancer specific genes (Table 5-3).

5.4 Discussion

Over the last few years the role of DNA methylation in cancer has been the subject of many studies. Recently our group has provided evidence for a new model of carcinogenesis. The predisposition of Polycomb Repressor Complex 2 (PRC2) targets to cancer-specific DNA hypermethylation suggests a 'crosstalk 'between PRC2 and de novo DNA methyltransferases in precursor cancer cells with a PRC2 distribution similar to that of ES cells. This 'crosstalk' may be initiated and/or facilitated by various environmental exposures, transgenerational inheritance, endocrine exposure, inflammation and by age. If a cell loses the potential to terminally differentiate as a consequence of irreversible CpG methylation, it will undergo prolonged exposure to environmental onslaught, and so, more likely to acquire those mutations and/or deletions necessary for carcinogenesis [280]. Based on these observations PCGT were chosen to be analysed. Additionally, it is generally accepted that methylation of gene promoters is associated with gene silencing. However, as accumulating evidence suggests that DNA methylation can occur at loci without an effect on gene expression we wanted to investigate the correlation between *NEUROD1* methylation and expression. Based on this analysis no association between DNA methylation and gene expression was found. These data are in agreement with previous reports suggesting that PCGT genes with tumour-specific promoter DNA methylation are not normally expressed in the epithelium of the tumour. It also provides further evidence on our previous published data that DNA methylation of PCGT genes in cancer may result in a residual stem-cell memory rather than a selective pressure for silencing these particular genes during carcinogenesis [280].

Therefore, based on the following three observations: 1) there is not a relationship between *NEUROD1* methylation and expression levels 2) lack of

NEUROD1 expression in the majority of the samples tested and 3) published findings that methylation of PCGT genes is a promising target for marker identification [305] [280] [279, 388], we wanted to further examine the predictive role of these genes in breast cancer and to examine whether they are affected by intra-tumour heterogeneity. MethyLight analysis of PCGT genes identified a constant panel of genes to be methylated in both central and peripheral tumour samples compared to controls, and non-parametric paired analysis indicated that there was no statistical significant difference between the methylation levels of the two zones within the cancer. This was true for both breast cancer specific genes and genes that were not specifically methylated in cancer. In order to investigate whether there is an association between the methylation changes observed in the two different tumour tissues, we performed correlation analysis showing 4 out of 13 genes to be positive associated.

NEUROD1 was one of the genes that did not show statistically significant differences in the methylation levels between tumour taken from the centre and the periphery suggesting it is homogeneously methylated within the tumour. This is an important finding as it further supports our previous report that *NEUROD1* methylation could be a good predictive marker in breast cancer as it is not affected by intra-tumour heterogeneity [305]. Moreover, this study provides further evidence for paired-like homeodomain transcription factor 2 (*PITX2*) which has been shown by Harbeck *et al*, to be a good biomarker for breast cancer hormone therapy treatment and, having performed analysis of several different tissue sections, has also shown a low variability in methylation measurements [393].

Through literature review, no other study was found to have examined breast cancer intra-tumour heterogeneity and its effect on DNA methylation changes. In contrast, there are two studies analysing expression modifications in correlation to intra-tumour heterogeneity. Both were carried on micro-dissected tumour cells rather than core biopsies. The first study by Aubele *et al* confirmed heterogeneity by comparative genomic hybridization [375]. The second study by Zhu *et al* described expression heterogeneity in sections that were obtained from morphologically dissimilar regions, one from the centre containing invasive breast tumourigenic cells, and the other from the periphery containing DCIS. The differences in the expression profile described in this study could however, be attributed to the different type of cells analysed i.e. comparing invasive to non-invasive cells [376]. Our finding that methylated PCGT genes provide reliable data irrespective of sampling topography, suggests that methylation analysis of these genes could hold great potential for improving breast cancer management. They could be useful for the early diagnosis of breast cancer predicting the biology of these tumours; refer to epigenetic treatment strategies, and finally provide suggestions that could have an important impact on the future of women's health. In addition, we have also demonstrated that the technology for methylation analysis can be easily applied in clinical routine as only a core biopsy would be required instead of purified cell population of cells.

Our results are in contrast to reports of DNA methylation changes of different candidate genes in other cancer types which have been shown to be affected by intra-tumour heterogeneity. In melanoma when methylation changes and expression status of suppressor genes were analysed, tissue taken from the centre of the tumour found to represent the whole tumour more accurately than

the tissue from the periphery [394, 395]. A more recent study of ovarian cancer suggested that both inter- and intra-tumour heterogeneity are allied with *NY-ESO-1* expression, which was correlated with promoter and global DNA-methylation alterations when micro-dissected cells were analysed [396]. Further studies are required before it can be determined whether these changes are cancer or gene specific.

In this study we have identified and confirmed in the validation set, 13 PCGTs that can predict breast cancer. This includes the first report of hypermethylation of the trans-membrane protein with EGF-like and two follistatin-like domains 2 (*TEMFF2*), the proenkephalin (*PENK*), glutamate decarboxylase-1 (*GAD1*) and cellular retinoic acid binding protein 1 (*GRABP1*) genes in breast cancer. Even though the role of *TMEFF2* gene methylation has been observed in other types of cancer such as colorectal, bladder [397, 398] and gastric adenocarcinomas [399] there are no reports for breast cancer. *PENK* gene expression has been shown to be down regulated in prostate and bladder cancer using, expression profiling [400]. Methylation of *GAD1* has not been previously shown to be associated with cancer and *CRABP1* methylation was only reported in association with colon [401] and ovarian [402] cancer. In contrast, there are reports of an association between methylation of *PITX2*, also an ER-targeted gene - the only ER-targeted gene analysed shown to be cancer specific -, and metallothionein 1A (*MT1A*) with breast carcinogenesis, further validating the data presented [393, 403].

Interestingly, 6 of the 13 methylated loci identified are genes belonging to the homeobox (HOX) domain. These genes are known to control normal development and differentiation of many multi-cellular organisms [404]. In humans

there are 39 HOX genes organised in four clusters – A, B, C and D being localised on chromosome 7, 17, 2 AND 12 respectively [389]. The last decade the role of the HOX domain in carcinogenesis has also been pointed. An example is *HOXB7* which has been implicated as an oncogene and is known to increase the expression of basic fibroblast growth factor (*bFGF*) in melanoma [405]. Indeed, previous data has confirmed that HOX gene cluster methylation is a common feature in cancer [390] including breast cancer [406, 407]. In the search for more specific and sensitive markers a recent study by Fiegl *et al*, has shown that methylation of *HOXA11* is strongly associated with the residual tumor after cytoreductive surgery and a good marker indicating poor prognosis in ovarian cancer [408]. A recent study though has suggests that methylation of the HOXA cluster may be a normal developmental and cell type specific process rather than a cancer specific mechanism [409] but their study subjects are not well characterised and the tumour tissue samples analysed with MeDIP should have been compared with normal tissue from controls in order to be able to have solid conclusions. It is worth mentioning that in this study despite the small sample size used the results are consistent and it has identified cancer specific genes characterized by high AUC values indicating a high sensitivity and specificity compared to studies that have used larger sample sizes [406].

Methylation of *HOXA9* was observed in morphological normal tissue adjacent to the breast tumour. There could be two possible explanations on why this is observed. The first explanation could be based on the theory of a cancer stem cell population. As it was discussed in the literature review, the last few years there is an increasing body of evidence suggesting that early epigenetic changes of stem cells may be the initiating events in carcinogenesis and evolve in normal

tissue in advance to tumour formation [115]. Methylation changes of non-tumour tissue adjacent to the tumour may represent an epigenetic disruption of progenitor cells which could lead to tumour formation through a stepwise process that could lead to further epigenetic changes. DNA methylation changes may lead to a polyclonal population of cells which have the potential for neoplastic changes. If this is true then the ideal target for cancer risk assessment and treatment would be to detect these pre-neoplastic epigenetic lesions before tumour formation. At the moment there are studies trying to identify and isolate cancer stem cells [410, 411] but in general these approaches have been hindered by technical difficulties.

The second possibility is that the methylation changes observed could represent a premalignant epigenetic lesion which is a mediator of a field defect in these tissues, although the origin of this field remains unknown [273]. Possibly these methylation changes around the tumour even though they are not responsible for causing transformation on their own, they could be permissive for the acquisition of additional epigenetic or genetic changes which could eventually lead to tumour formation. The first observation of this phenomenon was made in oral cancer [412]. This theory has been supported by subsequent studies carried out on colorectal cancer [272], head and neck squamous cell carcinoma, lung, esophagus, vulva, cervix, bladder, and skin and breast cancer [273]. Based on such observations it has been suggested that such information could prove valuable for risk assessment. The study on breast cancer identified *RASSF1A* to be methylated in four different zones of normal breast tissue in the ipsilateral and contralateral breasts of women with breast cancer [273]. In contrast to our study, they showed *RASSF1A* methylation to occur in both the tumour and the normal tissue adjacent

to the tumour with a higher degree of methylation seen in the tissue closer to the tumour [273].

In addition, the results showed *HOXA9* to be statistically significant less frequently methylated in the morphologically normal tissue adjacent to the tumour compared to the controls and the corresponding tumour tissue. This observation comes in agreement with another study by our group investigating methylation changes in white blood cells and whether they are predictive of breast cancer risk [163]. It has to be pointed that in this study *HOXA9* was also shown methylated in the tissue analysed from the controls. This could be mainly attributed to aging, in a recent study our group showed that stem cell PCGTs are far more likely to become methylated with age than non-targets [413]. Therefore, to avoid any bias in the study our study women were all postmenopausal and aged-matched. It needs to be pointed that the control samples used for the purposes of the study is not the ideal set of samples, the other option could have been (but not available) to use cancer free women undergoing mammoplasty reduction. The problem with this sample set though would have been the low number of progenitor cells for analysis.

For future work it would be interesting to investigate methylation profiles of metastatic specimens compared to matched primary tissues in order to examine whether DNA methylation of PCGT genes is homogeneous in breast cancer metastases. A recent study by Wu *et al* demonstrated that samples taken from a patient's primary breast carcinoma and their metastatic breast cancer are characterized by extensive expression heterogeneity. The study confirmed that ER and/or PR status characterising the primary cancer may be lost in the

metastatic carcinoma. This observation is important as the metastasis will not be hormone sensitive as its primary carcinoma resulting in resistance to the therapy. Interestingly the methylation signature of the primary tumour tissue compared to the metastatic specimen was similar, with the latter only exhibiting a higher intensity of methylation [414].

In addition, microdissection of epithelial and stromal cells from the morphological normal tissue adjacent to the tumour compared to the corresponding tumour and controls would be an interesting study to better understand which cells are triggering DNA methylation changes in the breast. Preliminary evidence suggests that DNA methylation in the epithelial genome could be directed by the neighbouring fibroblasts indicating that the breast cancer microenvironment may be the one inducing epigenetic changes [415]. Genetic changes in the morphologically normal cells adjacent to the tumour have been shown to be characterised by loss of heterozygosity, microsatellite and chromosome instability and gene mutations [416] with these alterations in the stroma not mimicking those in the epithelium playing a different and parallel role in carcinogenesis and tumor progression, probably by modifying some features specific to breast cancer [417]. A more recent study though has shown conflicting results [418].

Finally, it would be ideal to analyse tumour tissue and the corresponding plasma/serum from breast cancer patients but also plasma/serum collected before diagnosis from women with breast cancer and from controls to better understand the role of DNA methylation in breast cancer and their clinical value. A recent study has shown that aberrant promoter hypermethylation of *RASSFA1* n serum/plasma DNA may be common among high-risk women and may be present years before cancer diagnosis [419]. Another study has shown an

association between elevated levels of tenascin-W and the presence of cancer in both serum samples and the stoma tissue analysed in colorectal cancers [420].

This is a proof of principle study demonstrating for the first time that methylation of PCGT genes is unaffected by intra-tumour heterogeneity within a set of breast cancer samples. These findings suggest that methylation of specifically identified PCGT genes may present a more robust means with which to guide breast cancer management, particularly in instances when only small core biopsies are available for assessment. Further investigation of epigenetic intra-tumour heterogeneity within breast cancer, as well as other cancer types is necessary. In addition, more studies are needed in order to better understand field cancerisation in breast cancer and its value in risk prediction. Nonetheless evidence is beginning to accumulate in recognition of the potential of DNA methylation markers in cancer assessment and treatment.

6 SUMMARY AND FUTURE WORK

6.1 Introduction

The aim of the research presented in this PhD thesis is to identify and validate risk factors/biomarkers for breast cancer and add to the ongoing efforts to improve risk prediction in the general population. The project is continuing with analysis of serum DNA methylation changes using high-throughput technology to establish whether DNA methylation profiles will serve as a new tool to predict breast cancer risk. The plan is to correlate these changes with serum hormonal levels with final goal to build better breast cancer risk prediction models. A number of important findings have resulted from the first phase of the work which forms the basis of this thesis and are summarised in the following section.

6.2 Summary of the main findings

6.2.1 Which is the best way to identify breast cancer cases in the general population?

Identification of breast cancer cases within UKCTOCS was a lengthy process. When the study started it was decided to use not only notification obtained from the cancer registry and UKCTOCS FUQ but to validate this using a specific questionnaire on breast cancer (the BCQ) which was send to the treating physician of the women. The purpose was to confirm diagnosis and to collect further clinicopathological data. Identification of the women with breast cancer started in October 2006 and ended in February 2009 with the last up-date from ONS. It resulted in the initial identification of 2629 women with possible breast cancer. As soon as ethical approval was provided the BCQs were sent to the treating physician of the women. 1083 BCQs were returned. By comparing the three sources discrepancies were identified and the need to further investigate

which source is most accurate for breast cancer diagnosis was born. Literature review confirmed the absence of such a study with the majority of previous studies reporting on accuracy of cancer data by comparing self-reported data collected using questionnaires to CR records. It became apparent that this would be the largest study conducted in England and Wales examining the sensitivity of both self-reporting and CR to medical confirmation obtained in a form of questionnaire completed by the physicians treating the women diagnosed with breast cancer.

Main findings:

1) Decreased sensitivity of CR compared to FUQ is mostly due to delays in cancer registration with higher PPV characterising CR compared to self-reporting (FUQ).
2) Researchers could rely on national CR data as it had the lowest percentage of misclassifications.
3) Self-reported data is another good source but accuracy is to a certain extent dependent on factors such as age, education and family breast cancer history.
4) Only directly contacting physicians eliminated all discrepancies within our cohort which otherwise would have been misclassified.
5) Confirmation of the cancer diagnosis by checking medical notes as it was carried out using a questionnaire provided the most complete data. However it needs to be noted that this is labour intensive and 70% of responses were received.

Going through the analysis affords the degree of assurance that the suggested methodology is sufficiently robust to accurately identify breast cancer cases. This

is important as studies investigating cancer risk assessment are completely dependent on accurate data about cancer diagnosis.

6.2.2 Shedding light on the role of sex steroid hormones and their controllers -gonadotrophins- and examining for the first time serum bioactivity of sex steroid receptors in breast cancer

Sex steroid hormones are known to be involved in breast carcinogenesis. However, search of the literature revealed the absence of studies investigating sex steroid hormone levels in breast cancer patients at different time intervals before diagnosis and their role in risk prediction. The meta-analysis by Key et al [197] was the first report to examine whether differences exist in breast cancer risk less and more than two years before breast cancer diagnosis. Additionally, no study was found that had explored the role of gonadotrophins and breast cancer risk and there was only one study with adequate sample size that had investigated progesterone and breast cancer risk [202]. No reports were found investigating SB of steroid receptors in relation to breast cancer. With the recent report by our group demonstrating a 10 fold increase of breast cancer risk in women with ER-α and ER-β SB in the top quintile at the time of diagnosis [162], it was of great interest to investigate their value in samples collected up to five years before diagnosis both to asses them as markers and better understand their role in breast carcinogenesis. Finally, literature search showed the absence of studies investigating the joint effect of hormones in predicting breast cancer risk with all studies examining the effect of single hormones. The only study that had explored the combined effect of oestrogens and androgens had shown to increase breast cancer risk prediction but women had provided serum samples at the time of breast cancer diagnosis [201] and not prior to diagnosis years in

advance as in this study. Therefore, using nested case control study in UKCTOCS eligible cases and controls were identified to investigate and cover all the above missing links with the aim to better understand the role of sex steroid hormones and gonadotrophins in breast cancer risk prediction.

Main findings:

1) Less than 2 years before diagnosis, the main oestrogen associated with breast cancer risk is oestrone with increased levels being significantly associated with breast cancer risk.
2) The main androgen associated with breast cancer risk is testosterone with increased levels independent of time prior to diagnosis being associated with increased risk.
3) ER-α and ER-β SBs more than 2 years before diagnosis are associated with breast cancer with women having ER-α SB in top quintile 2 years before diagnosis having a two fold increased breast cancer risk.
4) Combination of hormones has a better breast cancer risk prediction power in comparison to single hormones.
5) Testosterone and FSH were shown to have a possible synergistic effect in breast carcinogenesis with a high risk predictive power independent of time to diagnosis.
6) SHBG and ER-β SB was demonstrated to have a possible synergistic effect in breast carcinogenesis but also to predict breast cancer more than 2 years before diagnosis with high significance.

6.2.3 What does the future hold of DNA methylation changes in breast cancer?

Epigenetics and especially DNA methylation changes hold a great promise in the future for clinical assessment of breast cancer. Researching the literature it was clear that other methods to assess the disease were lacking due to intra-tumour heterogeneity [375, 376]. Additionally, there was a huge interest in the environment around the tumour, with the majority of the studies on colon and breast cancer showing methylation changes in premalignant colorectal and breast tissue representing a field defect, perpetuating further neoplastic changes [272, 273]. In the meantime, our group among others was the first to provide evidence that stem cell PCGT genes are more likely to have cancer specific promoter DNA hypermethylation than non-PCGT genes [279, 280, 388]. Therefore, investigation of the role of PCGT methylation in breast cancer examining intra-tumour heterogeneity and epigenetic field defect was raised. With this study, it was shown for the first time that PCGT methylation changes were predictive of breast cancer and homogeneous across the tumour. This is an important finding as PCGT methylation changes could prove good candidate markers for serum analysis but before analysing serum it was fundamental to investigate whether such changes are representative of the entire tumour. The discovery of a cancer marker that is detected in both the serum and tumour tissue would be ideal – therefore, by looking into whether this marker is affected by intra-tumour heterogeneity is the first step towards further analysis.

Main findings:

1) Methylation of specific PCGT loci predicts the presence of breast cancer in core biopsy specimens.

2) 13/55 PCGT genes were shown to be cancer specific (p\0.05) with a ROC AUC of 0.7 (range 0.71–0.95) and with DNA methylation changes investigated predicting the presence of cancer in both tissues taken from the center and the periphery of the tumour.
3) Methylation of specific PCGT loci is unaffected by tumour-heterogeneity.
4) DNA methylation analysis of PCGT genes carries information independent from expression and could be used to assess core breast biopsies and ultimately guide patient management.

6.3 Future work

For future work we are planning to work on the following aspects (some of the suggested future work is already underway):

- Apart from the main role of oestrogens and androgens in breast cancer risk, it is of great interest to better understand the role progesterone in breast carcinogenesis. Previous studies have shown HRT use (synthetic progesterone derivatives are used) to be associated with an increased risk of breast cancer. In a recent study it has been demonstrated that *in vivo* administration of medroxyprogesterone acetate triggers massive induction of the key osteoclast differentiation factor Receptor Activator of NF-κB Ligand (RANKL) in mammary gland epithelial cells (data to be published in Nature by collaborators). RANKL is essential for the development and activation of osteoclasts. RANKL and its receptor RANK also control lymph node organogenesis, development of thymic medullary epithelial cells and, importantly, formation of a lactating mammary gland during pregnancy. Both RANKL and RANK expression have been observed in primary breast cancers in humans and breast cancer cells lines and

it has been proposed that the RANKL/RANK system can regulate bone metastases of epithelial tumours [421, 422]. Based on these observations, in a collaboration set up between Professor Martin Widschwendter with Professor Georg Schett and Professor Josef M. Penninger further studies are underway to investigate the role of RANK and RANKL and correlate it with progesterone levels in women described in chapter 4. Furthermore, PR SB will be investigated to better understand its association with breast cancer. Our collaborators in Bonn are trying to establish the assay. During my stay in Bonn, different clones produced by Guido Hasenbrink were tried under different conditions (this data was not shown) to check functionality of the assay but unfortunately the assay did not work. Once the PR SB assay is perfected the plan is to analyse PR SB in our cohort.

- To better understand the synergistic effect of FSH and testosterone in breast carcinogenesis further experiments will be carried out. Collaboration has already been set up with Professor Louis Dubeau to investigate the possible synergist effect of FSH and testosterone in mammary tissue in a mouse model. This will involve crossing two transgenic mouse lines expressing Cre recombinase under the control of a truncated form of the FSHR with a ROSA26R Cre reporter mouse and investigating the expression of FSHR in the breast tissue (stroma, epithelium and fat). If expression is confirmed mammary cells will be treated with FSH and IGF1 alone and in combination to investigate whether increased aromatization (CYP11A1, HSD3B1 and CYP19A1 mRNA levels measured) occurs by comparing them to cells that have not been treated.

- After investigating the homogeneity of PCGT methylation changes in breast tumour, it would be interesting to examine methylation profiles of metastatic specimens compared to matched primary tissues to study whether DNA methylation of PCGT genes are homogeneous in breast cancer metastases. Additionally, microdissected epithelial and stromal cells from the morphological normal tissue adjacent to the tumour compared to the corresponding tumour and controls would be an interesting study to better understand which cells are triggering DNA methylation changes in the breast.

- Analyses of tumour tissue and the corresponding plasma/serum from breast cancer patients but also plasma/serum collected before diagnosis from women with breast cancer and from controls to better understand the role of DNA methylation in breast cancer and their clinical value are essential. Over the last decade it has become clear that hypermethylation can be detected in tumour-derived DNA found in the serum and plasma of cancer patients. The far majority of studies have analysed serum/plasma in diseased individuals to either use this as an early detection marker or as a prognostic/predictive marker. None of the studies so far have addressed the question whether DNA in serum/plasma is able to predict predisposition to develop cancer years before onset of disease. A method to detect pre-neoplastic and/or early neoplastic change prior to tumour mass formation is needed to allow us to catch tumours early. Such an approach would also offer invaluable knowledge to add to current theories of carcinogenesis. During the PhD study new techniques were discovered for epigenome analysis caughting up with the demands of modern epigenetics. Further work will be carried out to using whole epigenome analysis to investigate whether serum DNA methylation changes could prove useful markers for risk

assessment. Our group has already started epigenotyping the 200 cases and 400 controls used to examine hormonal changes along with more cases that were further identified through the process described in chapter 3 using high-throughput technology (Illumina) to discover markers for breast cancer risk prediction. The work includes optimising DNA extraction from serum, analysing DNA methylation changes for more than 27,000 genes using Illumina technology in 350 cases and 400 controls, solving the statistical issues in the analysis of high-throughput DNA methylation data.

- Animal models have provided a lot of information linking the effects of steroid hormones on epigenome to cancer. Much less is available in human beings. Our group was one of the first to find in breast cancer, target genes of ER-α to be less methylated in ER positive cancers in comparison to ER negative ones [163]. One of the main questions to be answered through the planned work is whether long-term hormonal exposure alters the epigenome in non-neoplastic cells in human beings. In order to answer this question DNA methylation changes in serum samples described in chapter 4 will be correlated with the already measured sex steroid hormones and sex steroid receptor SB to better understand their effect on the epigenome and in breast carcinogenesis.

- Eventually combination of all the data will show whether a better risk prediction could be obtained that will eventually have an impact in women's life.

6.4 Conclusion

This thesis reported on all the important incremental steps made in achieving the aims described in chapter 1 within the timeframe of the three-year PhD. Briefly,

the aims of this project were to: 1) examine which is the best source of identifying breast cancer cases in the general population 2) investigate the association of sex steroids, gonadotrophins and novel assays of sex steroid hormone receptor SB in breast cancer 3) examine whether they can be combined to improve breast cancer risk assessment and investigate their synergistic effect 4) identify new DNA methylation markers that might add to such a strategy in the future, with an overall goal to improve breast cancer risk prediction. The findings of this research have shown that the most accurate source of information for breast cancer diagnosis involves combining CR and self-reporting data using the rule that both must concur if breast cancer is to be confirmed. The research has demonstrated that oestrone and testosterone are the most strongly associated oestrogens and androgens, respectively, with breast cancer risk along with SB of their receptors which proved to be an attractive alternative marker for risk assessment in postmenopausal women. By examining the best combination of hormones/SB for breast cancer risk prediction, testosterone and FSH were shown to have significant predictive power and a possible synergistic effect in breast carcinogenesis. Furthermore, DNA methylation changes were shown to be associated with breast cancer and most importantly to be homogeneous. This is important for future studies trying to identify markers for risk assessment. Further studies are required to assess the role of serum DNA methylation changes and breast cancer risk. Examining the effect of sex steroid hormones into the epigenome and combination of hormones and breast cancer specific genes in a model to better predict breast cancer risk requires exploration.

REFERENCES

1. Thompson A, Brennan K, Cox A, Gee J, Harcourt D, Harris A, et al. Evaluation of the current knowledge limitations in breast cancer research: a gap analysis. Breast Cancer Res. 2008;10(2):R26.
2. Baum M. Ramifications of screening for breast cancer: consent for screening. Bmj. 2006 Mar 25;332(7543):728.
3. Vaidya JS. Screening for disease--the good--the bad and the thoughtful. Int J Surg. 2005;3(2):107-12.
4. Gail MH, Brinton LA, Byar DP, Corle DK, Green SB, Schairer C, et al. Projecting individualized probabilities of developing breast cancer for white females who are being examined annually. J Natl Cancer Inst. 1989 Dec 20;81(24):1879-86.
5. Costantino JP, Gail MH, Pee D, Anderson S, Redmond CK, Benichou J, et al. Validation studies for models projecting the risk of invasive and total breast cancer incidence. J Natl Cancer Inst. 1999 Sep 15;91(18):1541-8.
6. Elmore JG, Fletcher SW. The risk of cancer risk prediction: "What is my risk of getting breast cancer"? J Natl Cancer Inst. 2006 Dec 6;98(23):1673-5.
7. Barlow WE, White E, Ballard-Barbash R, Vacek PM, Titus-Ernstoff L, Carney PA, et al. Prospective breast cancer risk prediction model for women undergoing screening mammography. J Natl Cancer Inst. 2006 Sep 6;98(17):1204-14.
8. Chen J, Pee D, Ayyagari R, Graubard B, Schairer C, Byrne C, et al. Projecting absolute invasive breast cancer risk in white women with a model that includes mammographic density. J Natl Cancer Inst. 2006 Sep 6;98(17):1215-26.
9. Santen RJ, Boyd NF, Chlebowski RT, Cummings S, Cuzick J, Dowsett M, et al. Critical assessment of new risk factors for breast cancer: considerations for development of an improved risk prediction model. Endocr Relat Cancer. 2007 Jun;14(2):169-87.
10. Levenson VV. Biomarkers for early detection of breast cancer: what, when, and where? Biochim Biophys Acta. 2007 Jun;1770(6):847-56.
11. Website SoftCRU. http://www.cancerresearchuk.org/.
12. http://www.statistics.gov.uk.
13. Russo J, Russo IH. Cellular basis of breast cancer susceptibility. Oncol Res. 1999;11(4):169-78.
14. Li CI, Anderson BO, Daling JR, Moe RE. Trends in incidence rates of invasive lobular and ductal breast carcinoma. Jama. 2003 Mar 19;289(11):1421-4.
15. Arpino G, Bardou VJ, Clark GM, Elledge RM. Infiltrating lobular carcinoma of the breast: tumor characteristics and clinical outcome. Breast Cancer Res. 2004;6(3):R149-56.
16. Li CI, Moe RE, Daling JR. Risk of mortality by histologic type of breast cancer among women aged 50 to 79 years. Arch Intern Med. 2003 Oct 13;163(18):2149-53.
17. Ernster VL, Ballard-Barbash R, Barlow WE, Zheng Y, Weaver DL, Cutter G, et al. Detection of ductal carcinoma in situ in women undergoing screening mammography. J Natl Cancer Inst. 2002 Oct 16;94(20):1546-54.
18. Kuerer HM, Albarracin CT, Yang WT, Cardiff RD, Brewster AM, Symmans WF, et al. Ductal carcinoma in situ: state of the science and roadmap to advance the field. J Clin Oncol. 2009 Jan 10;27(2):279-88.
19. Harris L, Fritsche H, Mennel R, Norton L, Ravdin P, Taube S, et al. American Society of Clinical Oncology 2007 update of recommendations for the

use of tumor markers in breast cancer. J Clin Oncol. 2007 Nov 20;25(33):5287-312.
20. Bast RC, Jr., Ravdin P, Hayes DF, Bates S, Fritsche H, Jr., Jessup JM, et al. 2000 update of recommendations for the use of tumor markers in breast and colorectal cancer: clinical practice guidelines of the American Society of Clinical Oncology. J Clin Oncol. 2001 Mar 15;19(6):1865-78.
21. Singletary SE, Allred C, Ashley P, Bassett LW, Berry D, Bland KI, et al. Revision of the American Joint Committee on Cancer staging system for breast cancer. J Clin Oncol. 2002 Sep 1;20(17):3628-36.
22. Cianfrocca M, Goldstein LJ. Prognostic and predictive factors in early-stage breast cancer. Oncologist. 2004;9(6):606-16.
23. Giordano SH, Buzdar AU, Smith TL, Kau SW, Yang Y, Hortobagyi GN. Is breast cancer survival improving? Cancer. 2004 Jan 1;100(1):44-52.
24. Elston CW, Ellis IO. Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up. Histopathology. 1991 Nov;19(5):403-10.
25. Meyer JS, Alvarez C, Milikowski C, Olson N, Russo I, Russo J, et al. Breast carcinoma malignancy grading by Bloom-Richardson system vs proliferation index: reproducibility of grade and advantages of proliferation index. Mod Pathol. 2005 Aug;18(8):1067-78.
26. Lapidus RG, Nass SJ, Davidson NE. The loss of estrogen and progesterone receptor gene expression in human breast cancer. J Mammary Gland Biol Neoplasia. 1998 Jan;3(1):85-94.
27. Menard S, Pupa SM, Campiglio M, Tagliabue E. Biologic and therapeutic role of HER2 in cancer. Oncogene. 2003 Sep 29;22(42):6570-8.
28. Veronesi U, Boyle P, Goldhirsch A, Orecchia R, Viale G. Breast cancer. Lancet. 2005 May 14-20;365(9472):1727-41.
29. Hoeijmakers JH. DNA damage, aging, and cancer. N Engl J Med. 2009 Oct 8;361(15):1475-85.
30. Fraga MF, Agrelo R, Esteller M. Cross-talk between aging and cancer: the epigenetic language. Ann N Y Acad Sci. 2007 Apr;1100:60-74.
31. Trichopoulos D. Hypothesis: does breast cancer originate in utero? Lancet. 1990 Apr 21;335(8695):939-40.
32. Park SK, Kang D, McGlynn KA, Garcia-Closas M, Kim Y, Yoo KY, et al. Intrauterine environments and breast cancer risk: meta-analysis and systematic review. Breast Cancer Res. 2008;10(1):R8.
33. Xue F, Michels KB. Intrauterine factors and risk of breast cancer: a systematic review and meta-analysis of current evidence. Lancet Oncol. 2007 Dec;8(12):1088-100.
34. McTiernan A, Kuniyuki A, Yasui Y, Bowen D, Burke W, Culver JB, et al. Comparisons of two breast cancer risk estimates in women with a family history of breast cancer. Cancer Epidemiol Biomarkers Prev. 2001 Apr;10(4):333-8.
35. Schwartz AG, King MC, Belle SH, Satariano WA, Swanson GM. Risk of breast cancer to relatives of young breast cancer patients. J Natl Cancer Inst. 1985 Oct;75(4):665-8.
36. Ottman R, Pike MC, King MC, Henderson BE. Practical guide for estimating risk for familial breast cancer. Lancet. 1983 Sep 3;2(8349):556-8.
37. Pharoah PD, Day NE, Duffy S, Easton DF, Ponder BA. Family history and the risk of breast cancer: a systematic review and meta-analysis. Int J Cancer. 1997 May 29;71(5):800-9.

38. Miki Y, Swensen J, Shattuck-Eidens D, Futreal PA, Harshman K, Tavtigian S, et al. A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. Science. 1994 Oct 7;266(5182):66-71.
39. Wooster R, Bignell G, Lancaster J, Swift S, Seal S, Mangion J, et al. Identification of the breast cancer susceptibility gene BRCA2. Nature. 1995 Dec 21-28;378(6559):789-92.
40. Antoniou A, Pharoah PD, Narod S, Risch HA, Eyfjord JE, Hopper JL, et al. Average risks of breast and ovarian cancer associated with BRCA1 or BRCA2 mutations detected in case Series unselected for family history: a combined analysis of 22 studies. Am J Hum Genet. 2003 May;72(5):1117-30.
41. Lichtenstein P, Holm NV, Verkasalo PK, Iliadou A, Kaprio J, Koskenvuo M, et al. Environmental and heritable factors in the causation of cancer--analyses of cohorts of twins from Sweden, Denmark, and Finland. N Engl J Med. 2000 Jul 13;343(2):78-85.
42. Pharoah PD, Antoniou AC, Easton DF, Ponder BA. Polygenes, risk prediction, and targeted prevention of breast cancer. N Engl J Med. 2008 Jun 26;358(26):2796-803.
43. Stratton MR, Rahman N. The emerging landscape of breast cancer susceptibility. Nat Genet. 2008 Jan;40(1):17-22.
44. Ye Z, Parry JM. The CYP17 MspA1 polymorphism and breast cancer risk: a meta-analysis. Mutagenesis. 2002 Mar;17(2):119-26.
45. Feigelson HS, Shames LS, Pike MC, Coetzee GA, Stanczyk FZ, Henderson BE. Cytochrome P450c17alpha gene (CYP17) polymorphism is associated with serum estrogen and progesterone concentrations. Cancer Res. 1998 Feb 15;58(4):585-7.
46. Haiman CA, Hankinson SE, Spiegelman D, Colditz GA, Willett WC, Speizer FE, et al. The relationship between a polymorphism in CYP17 with plasma hormone levels and breast cancer. Cancer Res. 1999 Mar 1;59(5):1015-20.
47. Garcia-Closas M, Herbstman J, Schiffman M, Glass A, Dorgan JF. Relationship between serum hormone concentrations, reproductive history, alcohol consumption and genetic polymorphisms in pre-menopausal women. Int J Cancer. 2002 Nov 10;102(2):172-8.
48. Tworoger SS, Chubak J, Aiello EJ, Ulrich CM, Atkinson C, Potter JD, et al. Association of CYP17, CYP19, CYP1B1, and COMT polymorphisms with serum and urinary sex hormone concentrations in postmenopausal women. Cancer Epidemiol Biomarkers Prev. 2004 Jan;13(1):94-101.
49. Easton DF, Pooley KA, Dunning AM, Pharoah PD, Thompson D, Ballinger DG, et al. Genome-wide association study identifies novel breast cancer susceptibility loci. Nature. 2007 Jun 28;447(7148):1087-93.
50. Hunter DJ, Kraft P, Jacobs KB, Cox DG, Yeager M, Hankinson SE, et al. A genome-wide association study identifies alleles in FGFR2 associated with risk of sporadic postmenopausal breast cancer. Nat Genet. 2007 Jul;39(7):870-4.
51. Garcia-Closas M, Chanock S. Genetic susceptibility loci for breast cancer by estrogen receptor status. Clin Cancer Res. 2008 Dec 15;14(24):8000-9.
52. Garcia-Closas M, Hall P, Nevanlinna H, Pooley K, Morrison J, Richesson DA, et al. Heterogeneity of breast cancer associations with five susceptibility loci by clinical and pathological characteristics. PLoS Genet. 2008 Apr;4(4):e1000054.
53. Stacey SN, Manolescu A, Sulem P, Rafnar T, Gudmundsson J, Gudjonsson SA, et al. Common variants on chromosomes 2q35 and 16q12 confer susceptibility to estrogen receptor-positive breast cancer. Nat Genet. 2007 Jul;39(7):865-9.

54. Stacey SN, Manolescu A, Sulem P, Thorlacius S, Gudjonsson SA, Jonsson GF, et al. Common variants on chromosome 5p12 confer susceptibility to estrogen receptor-positive breast cancer. Nat Genet. 2008 Jun;40(6):703-6.
55. Turnbull C, Ahmed S, Morrison J, Pernet D, Renwick A, Maranian M, et al. Genome-wide association study identifies five new breast cancer susceptibility loci. Nat Genet. 2010 May 9.
56. Vogel VG. Epidemiology, genetics, and risk evaluation of postmenopausal women at risk of breast cancer. Menopause. 2008 Jul-Aug;15(4 Suppl):782-9.
57. Page DL, Schuyler PA, Dupont WD, Jensen RA, Plummer WD, Jr., Simpson JF. Atypical lobular hyperplasia as a unilateral predictor of breast cancer risk: a retrospective cohort study. Lancet. 2003 Jan 11;361(9352):125-9.
58. Dupont WD, Parl FF, Hartmann WH, Brinton LA, Winfield AC, Worrell JA, et al. Breast cancer risk associated with proliferative breast disease and atypical hyperplasia. Cancer. 1993 Feb 15;71(4):1258-65.
59. Hartmann LC, Sellers TA, Frost MH, Lingle WL, Degnim AC, Ghosh K, et al. Benign breast disease and the risk of breast cancer. N Engl J Med. 2005 Jul 21;353(3):229-37.
60. Marshall LM, Hunter DJ, Connolly JL, Schnitt SJ, Byrne C, London SJ, et al. Risk of breast cancer associated with atypical hyperplasia of lobular and ductal types. Cancer Epidemiol Biomarkers Prev. 1997 May;6(5):297-301.
61. London SJ, Connolly JL, Schnitt SJ, Colditz GA. A prospective study of benign breast disease and the risk of breast cancer. Jama. 1992 Feb 19;267(7):941-4.
62. Boyd NF, Guo H, Martin LJ, Sun L, Stone J, Fishell E, et al. Mammographic density and the risk and detection of breast cancer. N Engl J Med. 2007 Jan 18;356(3):227-36.
63. Kerlikowske K, Ichikawa L, Miglioretti DL, Buist DS, Vacek PM, Smith-Bindman R, et al. Longitudinal measurement of clinical mammographic breast density to improve estimation of breast cancer risk. J Natl Cancer Inst. 2007 Mar 7;99(5):386-95.
64. Perry NM, Allgood PC, Milner SE, Mokbel K, Duffy SW. Mammographic breast density by area of residence: possible evidence of higher density in urban areas. Curr Med Res Opin. 2008 Feb;24(2):365-8.
65. Benson SR, Blue J, Judd K, Harman JE. Ultrasound is now better than mammography for the detection of invasive breast cancer. Am J Surg. 2004 Oct;188(4):381-5.
66. Gordon PB, Goldenberg SL. Malignant breast masses detected only by ultrasound. A retrospective review. Cancer. 1995 Aug 15;76(4):626-30.
67. Reduction in breast cancer mortality from organized service screening with mammography: 1. Further confirmation with extended data. Cancer Epidemiol Biomarkers Prev. 2006 Jan;15(1):45-51.
68. Humphrey LL, Helfand M, Chan BK, Woolf SH. Breast cancer screening: a summary of the evidence for the U.S. Preventive Services Task Force. Ann Intern Med. 2002 Sep 3;137(5 Part 1):347-60.
69. Tabar L, Fagerberg CJ, Gad A, Baldetorp L, Holmberg LH, Grontoft O, et al. Reduction in mortality from breast cancer after mass screening with mammography. Randomised trial from the Breast Cancer Screening Working Group of the Swedish National Board of Health and Welfare. Lancet. 1985 Apr 13;1(8433):829-32.
70. Shapiro S, Venet W, Strax P, Venet L, Roeser R. Ten- to fourteen-year effect of screening on breast cancer mortality. J Natl Cancer Inst. 1982 Aug;69(2):349-55.

71. Jorgensen KJ, Zahl PH, Gotzsche PC. Breast cancer mortality in organised mammography screening in Denmark: comparative study. Bmj.340:c1241.
72. Helvie MA, Chan HP, Adler DD, Boyd PG. Breast thickness in routine mammograms: effect on image quality and radiation dose. AJR Am J Roentgenol. 1994 Dec;163(6):1371-4.
73. Hartman AR, Daniel BL, Kurian AW, Mills MA, Nowels KW, Dirbas FM, et al. Breast magnetic resonance image screening and ductal lavage in women at high genetic risk for breast carcinoma. Cancer. 2004 Feb 1;100(3):479-89.
74. Leach MO, Boggis CR, Dixon AK, Easton DF, Eeles RA, Evans DG, et al. Screening with magnetic resonance imaging and mammography of a UK population at high familial risk of breast cancer: a prospective multicentre cohort study (MARIBS). Lancet. 2005 May 21-27;365(9473):1769-78.
75. Szabo BK, Aspelin P, Kristoffersen Wiberg M, Tot T, Bone B. Invasive breast cancer: correlation of dynamic MR features with prognostic factors. Eur Radiol. 2003 Nov;13(11):2425-35.
76. Wahl RL, Siegel BA, Coleman RE, Gatsonis CG. Prospective multicenter study of axillary nodal staging by positron emission tomography in breast cancer: a report of the staging breast cancer with PET Study Group. J Clin Oncol. 2004 Jan 15;22(2):277-85.
77. Pepe MS, Etzioni R, Feng Z, Potter JD, Thompson ML, Thornquist M, et al. Phases of biomarker development for early detection of cancer. J Natl Cancer Inst. 2001 Jul 18;93(14):1054-61.
78. Menon U, Gentry-Maharaj A, Ryan A, Sharma A, Burnell M, Hallett R, et al. Recruitment to multicentre trials--lessons from UKCTOCS: descriptive study. Bmj. 2008;337:a2079.
79. Gast MC, Schellens JH, Beijnen JH. Clinical proteomics in breast cancer: a review. Breast Cancer Res Treat. 2009 Jul;116(1):17-29.
80. Duffy MJ. Serum tumor markers in breast cancer: are they of clinical value? Clin Chem. 2006 Mar;52(3):345-51.
81. Wandall HH, Blixt O, Tarp MA, Pedersen JW, Bennett EP, Mandel U, et al. Cancer biomarkers defined by autoantibody signatures to aberrant O-glycopeptide epitopes. Cancer Res. 2010 Feb 15;70(4):1306-13.
82. Lane JA, Hamdy FC, Martin RM, Turner EL, Neal DE, Donovan JL. Latest results from the UK trials evaluating prostate cancer screening and treatment: The CAP and ProtecT studies. Eur J Cancer. 2010 Nov;46(17):3095-101.
83. Srinivas PR, Kramer BS, Srivastava S. Trends in biomarker research for cancer detection. Lancet Oncol. 2001 Nov;2(11):698-704.
84. Silva J, Silva JM, Garcia V, Garcia JM, Dominguez G, Bonilla F. RNA is more sensitive than DNA in identification of breast cancer patients bearing tumor nucleic acids in plasma. Genes Chromosomes Cancer. 2002 Dec;35(4):375-6.
85. Widschwendter M, Jones PA. DNA methylation and breast carcinogenesis. Oncogene. 2002 Aug 12;21(35):5462-82.
86. Widschwendter M, Menon U. Circulating methylated DNA: a new generation of tumor markers. Clin Cancer Res. 2006 Dec 15;12(24):7205-8.
87. Ullah MF, Aatif M. The footprints of cancer development: Cancer biomarkers. Cancer Treat Rev. 2009 May;35(3):193-200.
88. Hacker's Ba, editor. Gynecological Oncology: Lippincott Williams & Wilkins; 2008.
89. Smith I, Chua S. Medical treatment of early breast cancer. I: adjuvant treatment. Bmj. 2006 Jan 7;332(7532):34-7.
90. Smith I, Chua S. Medical treatment of early breast cancer. II: endocrine therapy. Bmj. 2006 Jan 14;332(7533):101-3.

91. Smith I, Chua S. Medical treatment of early breast cancer. III: chemotherapy. Bmj. 2006 Jan 21;332(7534):161-2.
92. Smith I, Chua S. Medical treatment of early breast cancer. IV: neoadjuvant treatment. Bmj. 2006 Jan 28;332(7535):223-4.
93. Clarke M, Collins R, Darby S, Davies C, Elphinstone P, Evans E, et al. Effects of radiotherapy and of differences in the extent of surgery for early breast cancer on local recurrence and 15-year survival: an overview of the randomised trials. Lancet. 2005 Dec 17;366(9503):2087-106.
94. Darby S, McGale P, Peto R, Granath F, Hall P, Ekbom A. Mortality from cardiovascular disease more than 10 years after radiotherapy for breast cancer: nationwide cohort study of 90 000 Swedish women. Bmj. 2003 Feb 1;326(7383):256-7.
95. Baum M, Vaidya JS. Targeted intra-operative radiotherapy-TARGIT for early breast cancer. Ann N Y Acad Sci. 2008 Sep;1138:132-5.
96. Vaidya JS, Tobias JS, Baum M, Keshtgar M, Joseph D, Wenz F, et al. Intraoperative radiotherapy for breast cancer. Lancet Oncol. 2004 Mar;5(3):165-73.
97. Vaidya JS, Joseph DJ, Tobias JS, Bulsara M, Wenz F, Saunders C, et al. Targeted intraoperative radiotherapy versus whole breast radiotherapy for breast cancer (TARGIT-A trial): an international, prospective, randomised, non-inferiority phase 3 trial. Lancet. 2010 Jul 10;376(9735):91-102.
98. Howell A, Dowsett M. Recent advances in endocrine therapy of breast cancer. Bmj. 1997 Oct 4;315(7112):863-6.
99. Ali S, Coombes RC. Endocrine-responsive breast cancer and strategies for combating resistance. Nat Rev Cancer. 2002 Feb;2(2):101-12.
100. Hayes DF. Tamoxifen: Dr. Jekyll and Mr. Hyde? J Natl Cancer Inst. 2004 Jun 16;96(12):895-7.
101. Dhingra K. Antiestrogens--tamoxifen, SERMs and beyond. Invest New Drugs. 1999;17(3):285-311.
102. Buzdar AU. Letrozole compared with tamoxifen as initial adjuvant therapy for breast cancer. J Clin Oncol. 2007 May 20;25(15):2147-8; author reply 8.
103. Forbes JF, Cuzick J, Buzdar A, Howell A, Tobias JS, Baum M. Effect of anastrozole and tamoxifen as adjuvant treatment for early-stage breast cancer: 100-month analysis of the ATAC trial. Lancet Oncol. 2008 Jan;9(1):45-53.
104. Dutta U, Pant K. Aromatase inhibitors: past, present and future in breast cancer therapy. Med Oncol. 2008;25(2):113-24.
105. Ovarian ablation or suppression in premenopausal early breast cancer: results from the international adjuvant breast cancer ovarian ablation or suppression randomized trial. J Natl Cancer Inst. 2007 Apr 4;99(7):516-25.
106. Friedrichs K, Holzel F, Janicke F. Combination of taxanes and anthracyclines in first-line chemotherapy of metastatic breast cancer: an interim report. Eur J Cancer. 2002 Sep;38(13):1730-8.
107. Spigel DR, Burstein HJ. HER2 overexpressing metastatic breast cancer. Curr Treat Options Oncol. 2002 Apr;3(2):163-74.
108. Slamon D, Pegram M. Rationale for trastuzumab (Herceptin) in adjuvant breast cancer trials. Semin Oncol. 2001 Feb;28(1 Suppl 3):13-9.
109. Engel RH, Kaklamani VG. HER2-positive breast cancer: current and future treatment strategies. Drugs. 2007;67(9):1329-41.
110. Bilancia D, Rosati G, Dinota A, Germano D, Romano R, Manzione L. Lapatinib in breast cancer. Ann Oncol. 2007 Jun;18 Suppl 6:vi26-30.
111. Wood ER, Truesdale AT, McDonald OB, Yuan D, Hassell A, Dickerson SH, et al. A unique structure for epidermal growth factor receptor bound to

GW572016 (Lapatinib): relationships among protein conformation, inhibitor off-rate, and receptor activity in tumor cells. Cancer Res. 2004 Sep 15;64(18):6652-9.
112. Yoo CB, Jones PA. Epigenetic therapy of cancer: past, present and future. Nat Rev Drug Discov. 2006 Jan;5(1):37-50.
113. Lyko F, Brown R. DNA methyltransferase inhibitors and the development of epigenetic cancer therapies. J Natl Cancer Inst. 2005 Oct 19;97(20):1498-506.
114. Egger G, Liang G, Aparicio A, Jones PA. Epigenetics in human disease and prospects for epigenetic therapy. Nature. 2004 May 27;429(6990):457-63.
115. Sharma S, Kelly TK, Jones PA. Epigenetics in Cancer. Carcinogenesis. 2009 Sep 13.
116. Allison Jones ML, Evangelia-Ourania Fourkala, Rebecca Kristeleit & Martin Widschwendter. Emerging promise of epigenetics and DNA methylation for the diagnosis and management of women's cancers. Epigenomics. 2010;2(1):9-38.
117. Wald NJ, Law MR, Duffy SW. Breast screening saves lives. Bmj. 2009;339:b2922.
118. Cummings SR, San Martin J, McClung MR, Siris ES, Eastell R, Reid IR, et al. Denosumab for prevention of fractures in postmenopausal women with osteoporosis. N Engl J Med. 2009 Aug 20;361(8):756-65.
119. Tyrer J, Duffy SW, Cuzick J. A breast cancer prediction model incorporating familial and personal risk factors. Stat Med. 2004 Apr 15;23(7):1111-30.
120. Parmigiani G, Berry D, Aguilar O. Determining carrier probabilities for breast cancer-susceptibility genes BRCA1 and BRCA2. Am J Hum Genet. 1998 Jan;62(1):145-58.
121. Chen WY, Colditz GA. Risk factors and hormone-receptor status: epidemiology, risk-prediction models and treatment implications for breast cancer. Nat Clin Pract Oncol. 2007 Jul;4(7):415-23.
122. Wacholder S, Hartge P, Prentice R, Garcia-Closas M, Feigelson HS, Diver WR, et al. Performance of common genetic variants in breast-cancer risk models. N Engl J Med. Mar 18;362(11):986-93.
123. Rosner B, Colditz GA, Iglehart JD, Hankinson SE. Risk prediction models with incomplete data with application to prediction of estrogen receptor-positive breast cancer: prospective data from the Nurses' Health Study. Breast Cancer Res. 2008;10(4):R55.
124. Chlebowski RT, Anderson GL, Lane DS, Aragaki AK, Rohan T, Yasmeen S, et al. Predicting risk of breast cancer in postmenopausal women by hormone receptor status. J Natl Cancer Inst. 2007 Nov 21;99(22):1695-705.
125. Clamp A, Danson S, Clemons M. Hormonal risk factors for breast cancer: identification, chemoprevention, and other intervention strategies. Lancet Oncol. 2002 Oct;3(10):611-9.
126. Cuzick J, Powles T, Veronesi U, Forbes J, Edwards R, Ashley S, et al. Overview of the main outcomes in breast-cancer prevention trials. Lancet. 2003 Jan 25;361(9354):296-300.
127. Fabian CJ, Kimler BF. Selective estrogen-receptor modulators for primary prevention of breast cancer. J Clin Oncol. 2005 Mar 10;23(8):1644-55.
128. Kalidas M, Brown P. Aromatase inhibitors for the treatment and prevention of breast cancer. Clin Breast Cancer. 2005 Apr;6(1):27-37.
129. Tata JR. One hundred years of hormones. EMBO Rep. 2005 Jun;6(6):490-6.
130. C.G.D.Brook NJM. Endocrinology. Blackwell Science Ltd. 2001 2001:98-119 and 59-71.

131. Keightley MC. Steroid receptor isoforms: exception or rules? Molecular and Cellular Endocrinology. 1998;137:1-5.
132. Kumar R, Thompson EB. The structure of the nuclear hormone receptors. Steroids. 1999 May;64(5):310-9.
133. Brosens JJ, Tullet J, Varshochi R, Lam EW. Steroid receptor action. Best Pract Res Clin Obstet Gynaecol. 2004 Apr;18(2):265-83.
134. Gronemeyer H, Gustafsson JA, Laudet V. Principles for modulation of the nuclear receptor superfamily. Nat Rev Drug Discov. 2004 Nov;3(11):950-64.
135. Jensen EV, DeSombre ER. Estrogen-receptor interaction. Science. 1973 Oct 12;182(108):126-34.
136. Kuiper GG, Enmark E, Pelto-Huikko M, Nilsson S, Gustafsson JA. Cloning of a novel receptor expressed in rat prostate and ovary. Proc Natl Acad Sci U S A. 1996 Jun 11;93(12):5925-30.
137. Menasce LP, White GR, Harrison CJ, Boyle JM. Localization of the estrogen receptor locus (ESR) to chromosome 6q25.1 by FISH and a simple post-FISH banding technique. Genomics. 1993 Jul;17(1):263-5.
138. Enmark E, Pelto-Huikko M, Grandien K, Lagercrantz S, Lagercrantz J, Fried G, et al. Human estrogen receptor beta-gene structure, chromosomal localization, and expression pattern. J Clin Endocrinol Metab. 1997 Dec;82(12):4258-65.
139. Matthews J, Gustafsson JA. Estrogen signaling: a subtle balance between ER alpha and ER beta. Mol Interv. 2003 Aug;3(5):281-92.
140. Delaunay F, Pettersson K, Tujague M, Gustafsson JA. Functional differences between the amino-terminal domains of estrogen receptors alpha and beta. Mol Pharmacol. 2000 Sep;58(3):584-90.
141. Kuiper GG, Carlsson B, Grandien K, Enmark E, Haggblad J, Nilsson S, et al. Comparison of the ligand binding specificity and transcript tissue distribution of estrogen receptors alpha and beta. Endocrinology. 1997 Mar;138(3):863-70.
142. Wilson CM, Griffin JE, Wilson JD, Marcelli M, Zoppi S, McPhaul MJ. Immunoreactive androgen receptor expression in subjects with androgen resistance. J Clin Endocrinol Metab. 1992 Dec;75(6):1474-8.
143. Wilson CM, McPhaul MJ. A and B forms of the androgen receptor are expressed in a variety of human tissues. Mol Cell Endocrinol. 1996 Jun 18;120(1):51-7.
144. Cobinet J, Poujol, N., Soultan C. Molecular action of androgens Molecular and Cellular Endocrinology. 2002;198:15-24.
145. Misrahi M, Venencie PY, Saugier-Veber P, Sar S, Dessen P, Milgrom E. Structure of the human progesterone receptor gene. Biochim Biophys Acta. 1993 Nov 16;1216(2):289-92.
146. Pieber D, Allport VC, Bennett PR. Progesterone receptor isoform A inhibits isoform B-mediated transactivation in human amnion. Eur J Pharmacol. 2001 Sep 7;427(1):7-11.
147. Li X, Lonard DM, O'Malley BW. A contemporary understanding of progesterone receptor function. Mech Ageing Dev. 2004 Oct-Nov;125(10-11):669-78.
148. Simoni M, Gromoll J, Nieschlag E. The follicle-stimulating hormone receptor: biochemistry, molecular biology, physiology, and pathophysiology. Endocr Rev. 1997 Dec;18(6):739-73.
149. Chatterjee SK, Zetter BR. Cancer biomarkers: knowing the present and predicting the future. Future Oncol. 2005 Feb;1(1):37-50.
150. Matern D, Magera MJ. Mass spectrometry methods for metabolic and health assessment. J Nutr. 2001 May;131(5):1615S-20S.

151. Dorgan JF, Fears TR, McMahon RP, Aronson Friedman L, Patterson BH, Greenhut SF. Measurement of steroid sex hormones in serum: a comparison of radioimmunoassay and mass spectrometry. Steroids. 2002 Mar;67(3-4):151-8.
152. Khosla S, Amin S, Singh RJ, Atkinson EJ, Melton LJ, 3rd, Riggs BL. Comparison of sex steroid measurements in men by immunoassay versus mass spectroscopy and relationships with cortical and trabecular volumetric bone mineral density. Osteoporos Int. 2008 Oct;19(10):1465-71.
153. Couzinet B, Meduri G, Lecce MG, Young J, Brailly S, Loosfelt H, et al. The postmenopausal ovary is not a major androgen-producing gland. J Clin Endocrinol Metab. 2001 Oct;86(10):5060-6.
154. Fritz MA, Speroff L. The endocrinology of the menstrual cycle: the interaction of folliculogenesis and neuroendocrine mechanisms. Fertil Steril. 1982 Nov;38(5):509-29.
155. Nilsson S, Makela S, Treuter E, Tujague M, Thomsen J, Andersson G, et al. Mechanisms of estrogen action. Physiol Rev. 2001 Oct;81(4):1535-65.
156. Hormones and cancer: 90 years after Beatson. Cancer Surv. 1986;5(3):435-687.
157. Pike MC, Krailo MD, Henderson BE, Casagrande JT, Hoel DG. 'Hormonal' risk factors, 'breast tissue age' and the age-incidence of breast cancer. Nature. 1983 Jun 30;303(5920):767-70.
158. Santen R, Cavalieri E, Rogan E, Russo J, Guttenplan J, Ingle J, et al. Estrogen mediation of breast tumor formation involves estrogen receptor-dependent, as well as independent, genotoxic effects. Ann N Y Acad Sci. 2009 Feb;1155:132-40.
159. Pike MC, Spicer DV, Dahmoush L, Press MF. Estrogens, progestogens, normal breast cell proliferation, and breast cancer risk. Epidemiol Rev. 1993;15(1):17-35.
160. Conzen SD. Minireview: nuclear receptors and breast cancer. Mol Endocrinol. 2008 Oct;22(10):2215-28.
161. Matthews J, Wihlen B, Tujague M, Wan J, Strom A, Gustafsson JA. Estrogen receptor (ER) beta modulates ERalpha-mediated transcriptional activation by altering the recruitment of c-Fos and c-Jun to estrogen-responsive promoters. Mol Endocrinol. 2006 Mar;20(3):534-43.
162. Widschwendter M, Lichtenberg-Frate H, Hasenbrink G, Schwarzer S, Dawnay A, Lam A, et al. Serum oestrogen receptor alpha and beta bioactivity are independently associated with breast cancer: a proof of principle study. Br J Cancer. 2009 Jul 7;101(1):160-5.
163. Widschwendter M, Apostolidou S, Raum E, Rothenbacher D, Fiegl H, Menon U, et al. Epigenotyping in peripheral blood cell DNA and breast cancer risk: a proof of principle study. PLoS One. 2008;3(7):e2656.
164. von Schoultz B. Androgens and the breast. Maturitas. 2007 May 20;57(1):47-9.
165. Lanari C, Molinolo AA. Progesterone receptors--animal models and cell signalling in breast cancer. Diverse activation pathways for the progesterone receptor: possible implications for breast biology and cancer. Breast Cancer Res. 2002;4(6):240-3.
166. Ross RK, Paganini-Hill A, Wan PC, Pike MC. Effect of hormone replacement therapy on breast cancer risk: estrogen versus estrogen plus progestin. J Natl Cancer Inst. 2000 Feb 16;92(4):328-32.
167. Tanaka Y, Kuwabara K, Okazaki T, Fujita T, Oizumi I, Kaiho S, et al. Gonadotropins stimulate growth of MCF-7 human breast cancer cells by

promoting intracellular conversion of adrenal androgens to estrogens. Oncology. 2000;59 Suppl 1:19-23.
168. Key TJ, Pike MC. The role of oestrogens and progestagens in the epidemiology and prevention of breast cancer. Eur J Cancer Clin Oncol. 1988 Jan;24(1):29-43.
169. Newcomb PA, Longnecker MP, Storer BE, Mittendorf R, Baron J, Clapp RW, et al. Recent oral contraceptive use and risk of breast cancer (United States). Cancer Causes Control. 1996 Sep;7(5):525-32.
170. Breast cancer and hormonal contraceptives: further results. Collaborative Group on Hormonal Factors in Breast Cancer. Contraception. 1996 Sep;54(3 Suppl):1S-106S.
171. Althuis MD, Brogan DR, Coates RJ, Daling JR, Gammon MD, Malone KE, et al. Hormonal content and potency of oral contraceptives and breast cancer risk among young women. Br J Cancer. 2003 Jan 13;88(1):50-7.
172. Cuzick J. Hormone replacement therapy and the risk of breast cancer. Eur J Cancer. 2008 Nov;44(16):2344-9.
173. Roy JA, Sawka CA, Pritchard KI. Hormone replacement therapy in women with breast cancer. Do the risks outweigh the benefits? J Clin Oncol. 1996 Mar;14(3):997-1006.
174. Steinberg KK, Thacker SB, Smith SJ, Stroup DF, Zack MM, Flanders WD, et al. A meta-analysis of the effect of estrogen replacement therapy on the risk of breast cancer. Jama. 1991 Apr 17;265(15):1985-90.
175. Colditz GA, Egan KM, Stampfer MJ. Hormone replacement therapy and risk of breast cancer: results from epidemiologic studies. Am J Obstet Gynecol. 1993 May;168(5):1473-80.
176. Breast cancer and hormone replacement therapy: collaborative reanalysis of data from 51 epidemiological studies of 52,705 women with breast cancer and 108,411 women without breast cancer. Collaborative Group on Hormonal Factors in Breast Cancer. Lancet. 1997 Oct 11;350(9084):1047-59.
177. Rossouw JE, Anderson GL, Prentice RL, LaCroix AZ, Kooperberg C, Stefanick ML, et al. Risks and benefits of estrogen plus progestin in healthy postmenopausal women: principal results From the Women's Health Initiative randomized controlled trial. Jama. 2002 Jul 17;288(3):321-33.
178. Shumaker SA, Legault C, Rapp SR, Thal L, Wallace RB, Ockene JK, et al. Estrogen plus progestin and the incidence of dementia and mild cognitive impairment in postmenopausal women: the Women's Health Initiative Memory Study: a randomized controlled trial. Jama. 2003 May 28;289(20):2651-62.
179. Beral V. Breast cancer and hormone-replacement therapy in the Million Women Study. Lancet. 2003 Aug 9;362(9382):419-27.
180. Velie EM, Nechuta S, Osuch JR. Lifetime reproductive and anthropometric risk factors for breast cancer in postmenopausal women. Breast Dis. 2005;24:17-35.
181. Garcia-Closas M, Brinton LA, Lissowska J, Chatterjee N, Peplonska B, Anderson WF, et al. Established breast cancer risk factors by clinically important tumour characteristics. Br J Cancer. 2006 Jul 3;95(1):123-9.
182. Breast cancer and breastfeeding: collaborative reanalysis of individual data from 47 epidemiological studies in 30 countries, including 50302 women with breast cancer and 96973 women without the disease. Lancet. 2002 Jul 20;360(9328):187-95.
183. Ma H, Bernstein L, Pike MC, Ursin G. Reproductive factors and breast cancer risk according to joint estrogen and progesterone receptor status: a meta-analysis of epidemiological studies. Breast Cancer Res. 2006;8(4):R43.

184. Reeves GK, Kan SW, Key T, Tjonneland A, Olsen A, Overvad K, et al. Breast cancer risk in relation to abortion: Results from the EPIC study. Int J Cancer. 2006 Oct 1;119(7):1741-5.
185. Friedenreich CM. Review of anthropometric factors and breast cancer risk. Eur J Cancer Prev. 2001 Feb;10(1):15-32.
186. Lahmann PH, Hoffmann K, Allen N, van Gils CH, Khaw KT, Tehard B, et al. Body size and breast cancer risk: findings from the European Prospective Investigation into Cancer And Nutrition (EPIC). Int J Cancer. 2004 Sep 20;111(5):762-71.
187. Rinaldi S, Key TJ, Peeters PH, Lahmann PH, Lukanova A, Dossus L, et al. Anthropometric measures, endogenous sex steroids and breast cancer risk in postmenopausal women: a study within the EPIC cohort. Int J Cancer. 2006 Jun 1;118(11):2832-9.
188. Borgquist S, Jirstrom K, Anagnostaki L, Manjer J, Landberg G. Anthropometric factors in relation to different tumor biological subgroups of postmenopausal breast cancer. Int J Cancer. 2009 Jan 15;124(2):402-11.
189. Missmer SA, Smith-Warner SA, Spiegelman D, Yaun SS, Adami HO, Beeson WL, et al. Meat and dairy food consumption and breast cancer: a pooled analysis of cohort studies. Int J Epidemiol. 2002 Feb;31(1):78-85.
190. Smith-Warner SA, Spiegelman D, Adami HO, Beeson WL, van den Brandt PA, Folsom AR, et al. Types of dietary fat and breast cancer: a pooled analysis of cohort studies. Int J Cancer. 2001 Jun 1;92(5):767-74.
191. Mense SM, Hei TK, Ganju RK, Bhat HK. Phytoestrogens and breast cancer prevention: possible mechanisms of action. Environ Health Perspect. 2008 Apr;116(4):426-33.
192. Seitz HK, Stickel F. Molecular mechanisms of alcohol-mediated carcinogenesis. Nat Rev Cancer. 2007 Aug;7(8):599-612.
193. Smith-Warner SA, Spiegelman D, Yaun SS, van den Brandt PA, Folsom AR, Goldbohm RA, et al. Alcohol and breast cancer in women: a pooled analysis of cohort studies. Jama. 1998 Feb 18;279(7):535-40.
194. Conlon MS, Johnson KC, Bewick MA, Lafrenie RM, Donner A. Smoking (active and passive), N-acetyltransferase 2, and risk of breast cancer. Cancer Epidemiol. Feb 28.
195. Trichopoulos D, Katsouyanni K. Oral contraceptives, tobacco smoking, and breast cancer risk. Lancet. 1989 Jul 15;2(8655):158.
196. Verkasalo PK, Thomas HV, Appleby PN, Davey GK, Key TJ. Circulating levels of sex hormones and their relation to risk factors for breast cancer: a cross-sectional study in 1092 pre- and postmenopausal women (United Kingdom). Cancer Causes Control. 2001 Jan;12(1):47-59.
197. Key T, Appleby P, Barnes I, Reeves G. Endogenous sex hormones and breast cancer in postmenopausal women: reanalysis of nine prospective studies. J Natl Cancer Inst. 2002 Apr 17;94(8):606-16.
198. Kaaks R, Rinaldi S, Key TJ, Berrino F, Peeters PH, Biessy C, et al. Postmenopausal serum androgens, oestrogens and breast cancer risk: the European prospective investigation into cancer and nutrition. Endocr Relat Cancer. 2005 Dec;12(4):1071-82.
199. Zeleniuch-Jacquotte A, Shore RE, Koenig KL, Akhmedkhanov A, Afanasyeva Y, Kato I, et al. Postmenopausal levels of oestrogen, androgen, and SHBG and breast cancer: long-term results of a prospective study. Br J Cancer. 2004 Jan 12;90(1):153-9.

200. Zeleniuch-Jacquotte A, Gu Y, Shore RE, Koenig KL, Arslan AA, Kato I, et al. Postmenopausal levels of sex hormones and risk of breast carcinoma in situ: results of a prospective study. Int J Cancer. 2005 Mar 20;114(2):323-7.
201. Adly L, Hill D, Sherman ME, Sturgeon SR, Fears T, Mies C, et al. Serum concentrations of estrogens, sex hormone-binding globulin, and androgens and risk of breast cancer in postmenopausal women. Int J Cancer. 2006 Nov 15;119(10):2402-7.
202. Missmer SA, Eliassen AH, Barbieri RL, Hankinson SE. Endogenous estrogen, androgen, and progesterone concentrations and breast cancer risk among postmenopausal women. J Natl Cancer Inst. 2004 Dec 15;96(24):1856-65.
203. Helzlsouer KJ, Alberg AJ, Bush TL, Longcope C, Gordon GB, Comstock GW. A prospective study of endogenous hormones and breast cancer. Cancer Detect Prev. 1994;18(2):79-85.
204. Sieri S, Krogh V, Bolelli G, Abagnato CA, Grioni S, Pala V, et al. Sex hormone levels, breast cancer risk, and cancer receptor status in postmenopausal women: the ORDET cohort. Cancer Epidemiol Biomarkers Prev. 2009 Jan;18(1):169-76.
205. Baglietto L, Severi G, English DR, Krishnan K, Hopper JL, McLean C, et al. Circulating steroid hormone levels and risk of breast cancer for postmenopausal women. Cancer Epidemiol Biomarkers Prev. Feb;19(2):492-502.
206. Beattie MS, Costantino JP, Cummings SR, Wickerham DL, Vogel VG, Dowsett M, et al. Endogenous sex hormones, breast cancer risk, and tamoxifen response: an ancillary study in the NSABP Breast Cancer Prevention Trial (P-1). J Natl Cancer Inst. 2006 Jan 18;98(2):110-5.
207. Eliassen AH, Missmer SA, Tworoger SS, Hankinson SE. Endogenous steroid hormone concentrations and risk of breast cancer: does the association vary by a woman's predicted breast cancer risk? J Clin Oncol. 2006 Apr 20;24(12):1823-30.
208. Tworoger SS, Missmer SA, Barbieri RL, Willett WC, Colditz GA, Hankinson SE. Plasma sex hormone concentrations and subsequent risk of breast cancer among women using postmenopausal hormones. J Natl Cancer Inst. 2005 Apr 20;97(8):595-602.
209. Kaaks R, Berrino F, Key T, Rinaldi S, Dossus L, Biessy C, et al. Serum sex steroids in premenopausal women and breast cancer risk within the European Prospective Investigation into Cancer and Nutrition (EPIC). J Natl Cancer Inst. 2005 May 18;97(10):755-65.
210. Micheli A, Muti P, Secreto G, Krogh V, Meneghini E, Venturelli E, et al. Endogenous sex hormones and subsequent breast cancer in premenopausal women. Int J Cancer. 2004 Nov 1;112(2):312-8.
211. Manjer J, Johansson R, Berglund G, Janzon L, Kaaks R, Agren A, et al. Postmenopausal breast cancer risk in relation to sex steroid hormones, prolactin and SHBG (Sweden). Cancer Causes Control. 2003 Sep;14(7):599-607.
212. Toniolo PG, Levitz M, Zeleniuch-Jacquotte A, Banerjee S, Koenig KL, Shore RE, et al. A prospective study of endogenous estrogens and breast cancer in postmenopausal women. J Natl Cancer Inst. 1995 Feb 1;87(3):190-7.
213. Milliken EL, Ameduri RK, Landis MD, Behrooz A, Abdul-Karim FW, Keri RA. Ovarian hyperstimulation by LH leads to mammary gland hyperplasia and cancer predisposition in transgenic mice. Endocrinology. 2002 Sep;143(9):3671-80.
214. Syed V, Ulinski G, Mok SC, Yiu GK, Ho SM. Expression of gonadotropin receptor and growth responses to key reproductive hormones in normal and

malignant human ovarian surface epithelial cells. Cancer Res. 2001 Sep 15;61(18):6768-76.
215. McSorley MA, Alberg AJ, Allen DS, Allen NE, Brinton LA, Dorgan JF, et al. Prediagnostic circulating follicle stimulating hormone concentrations and ovarian cancer risk. Int J Cancer. 2009 Aug 1;125(3):674-9.
216. Vaidya JS. Re: Declines in invasive breast cancer and use of postmenopausal hormone therapy in a screening mammography population. J Natl Cancer Inst. 2008 Apr 16;100(8):598-9; author reply 9.
217. Key TJ, Wang DY, Brown JB, Hermon C, Allen DS, Moore JW, et al. A prospective study of urinary oestrogen excretion and breast cancer risk. Br J Cancer. 1996 Jun;73(12):1615-9.
218. Dorgan JF, Longcope C, Stephenson HE, Jr., Falk RT, Miller R, Franz C, et al. Relation of prediagnostic serum estrogen and androgen levels to breast cancer risk. Cancer Epidemiol Biomarkers Prev. 1996 Jul;5(7):533-9.
219. Berrino F, Muti P, Micheli A, Bolelli G, Krogh V, Sciajno R, et al. Serum sex hormone levels after menopause and subsequent breast cancer. J Natl Cancer Inst. 1996 Mar 6;88(5):291-6.
220. Dorgan JF, Longcope C, Stephenson HE, Jr., Falk RT, Miller R, Franz C, et al. Serum sex hormone levels are related to breast cancer risk in postmenopausal women. Environ Health Perspect. 1997 Apr;105 Suppl 3:583-5.
221. Dorgan JF, Stanczyk FZ, Longcope C, Stephenson HE, Jr., Chang L, Miller R, et al. Relationship of serum dehydroepiandrosterone (DHEA), DHEA sulfate, and 5-androstene-3 beta, 17 beta-diol to risk of breast cancer in postmenopausal women. Cancer Epidemiol Biomarkers Prev. 1997 Mar;6(3):177-81.
222. Thomas H, Key T, Allen D, Moore J, Dowsett M, Fentiman I, et al. A prospective study of endogenous serum hormone concentrations and breast cancer risk in postmenopausal women. Am J Epidemiol. 1997 Jun 1;145(11):291-.
223. Zeleniuch-Jacquotte A, Bruning PF, Bonfrer JM, Koenig KL, Shore RE, Kim MY, et al. Relation of serum levels of testosterone and dehydroepiandrosterone sulfate to risk of breast cancer in postmenopausal women. Am J Epidemiol. 1997 Jun 1;145(11):1030-8.
224. Hankinson SE, Willett WC, Manson JE, Colditz GA, Hunter DJ, Spiegelman D, et al. Plasma sex steroid hormone levels and risk of breast cancer in postmenopausal women. J Natl Cancer Inst. 1998 Sep 2;90(17):1292-9.
225. Cauley JA, Lucas FL, Kuller LH, Stone K, Browner W, Cummings SR. Elevated serum estradiol and testosterone concentrations are associated with a high risk for breast cancer. Study of Osteoporotic Fractures Research Group. Ann Intern Med. 1999 Feb 16;130(4 Pt 1):270-7.
226. Kabuto M, Akiba S, Stevens RG, Neriishi K, Land CE. A prospective study of estradiol and breast cancer in Japanese women. Cancer Epidemiol Biomarkers Prev. 2000 Jun;9(6):575-9.
227. Onland-Moret NC, Kaaks R, van Noord PA, Rinaldi S, Key T, Grobbee DE, et al. Urinary endogenous sex hormone levels and the risk of postmenopausal breast cancer. Br J Cancer. 2003 May 6;88(9):1394-9.
228. Lamar CA, Dorgan JF, Longcope C, Stanczyk FZ, Falk RT, Stephenson HE, Jr. Serum sex hormones and breast cancer risk factors in postmenopausal women. Cancer Epidemiol Biomarkers Prev. 2003 Apr;12(4):380-3.
229. Feinberg AP, Tycko B. The history of cancer epigenetics. Nat Rev Cancer. 2004 Feb;4(2):143-53.

230. Euhus DM, Bu D, Milchgrub S, Xie XJ, Bian A, Leitch AM, et al. DNA methylation in benign breast epithelium in relation to age and breast cancer risk. Cancer Epidemiol Biomarkers Prev. 2008 May;17(5):1051-9.
231. Bird A. DNA methylation patterns and epigenetic memory. Genes Dev. 2002 Jan 1;16(1):6-21.
232. Bird AP, Wolffe AP. Methylation-induced repression--belts, braces, and chromatin. Cell. 1999 Nov 24;99(5):451-4.
233. Hendrich B, Bird A. Mammalian methyltransferases and methyl-CpG-binding domains: proteins involved in DNA methylation. Curr Top Microbiol Immunol. 2000;249:55-74.
234. Esteller M. Epigenetic gene silencing in cancer: the DNA hypermethylome. Hum Mol Genet. 2007 Apr 15;16 Spec No 1:R50-9.
235. Kouzarides T. Chromatin modifications and their function. Cell. 2007 Feb 23;128(4):693-705.
236. Miranda TB, Jones PA. DNA methylation: the nuts and bolts of repression. J Cell Physiol. 2007 Nov;213(2):384-90.
237. Ohm JE, Baylin SB. Stem cell chromatin patterns: an instructive mechanism for DNA hypermethylation? Cell Cycle. 2007 May 2;6(9):1040-3.
238. Gronbaek K, Hother C, Jones PA. Epigenetic changes in cancer. Apmis. 2007 Oct;115(10):1039-59.
239. Jones PA. DNA methylation and cancer. Oncogene. 2002 Aug 12;21(35):5358-60.
240. Costello JF, Plass C. Methylation matters. J Med Genet. 2001 May;38(5):285-303.
241. Bird AP, Southern EM. Use of restriction enzymes to study eukaryotic DNA methylation: I. The methylation pattern in ribosomal DNA from Xenopus laevis. J Mol Biol. 1978 Jan 5;118(1):27-47.
242. Bedford MT, van Helden PD. A method to analyze allele-specific methylation. Biotechniques. 1990 Dec;9(6):744-8.
243. Hatada I, Hayashizaki Y, Hirotsune S, Komatsubara H, Mukai T. A genomic scanning method for higher organisms using restriction sites as landmarks. Proc Natl Acad Sci U S A. 1991 Nov 1;88(21):9523-7.
244. Huang TH, Perry MR, Laux DE. Methylation profiling of CpG islands in human breast cancer cells. Hum Mol Genet. 1999 Mar;8(3):459-70.
245. Rauch T, Pfeifer GP. Methylated-CpG island recovery assay: a new technique for the rapid detection of methylated-CpG islands in cancer. Lab Invest. 2005 Sep;85(9):1172-80.
246. Tommasi S, Karm DL, Wu X, Yen Y, Pfeifer GP. Methylation of homeobox genes is a frequent and early epigenetic event in breast cancer. Breast Cancer Res. 2009;11(1):R14.
247. Weber M, Davies JJ, Wittig D, Oakeley EJ, Haase M, Lam WL, et al. Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells. Nat Genet. 2005 Aug;37(8):853-62.
248. Frommer M, McDonald LE, Millar DS, Collis CM, Watt F, Grigg GW, et al. A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. Proc Natl Acad Sci U S A. 1992 Mar 1;89(5):1827-31.
249. Herman JG, Graff JR, Myohanen S, Nelkin BD, Baylin SB. Methylation-specific PCR: a novel PCR assay for methylation status of CpG islands. Proc Natl Acad Sci U S A. 1996 Sep 3;93(18):9821-6.

250. Eads CA, Danenberg KD, Kawakami K, Saltz LB, Blake C, Shibata D, et al. MethyLight: a high-throughput assay to measure DNA methylation. Nucleic Acids Res. 2000 Apr 15;28(8):E32.
251. Xiong Z, Laird PW. COBRA: a sensitive and quantitative DNA methylation assay. Nucleic Acids Res. 1997 Jun 15;25(12):2532-4.
252. Ronaghi M, Uhlen M, Nyren P. A sequencing method based on real-time pyrophosphate. Science. 1998 Jul 17;281(5375):363, 5.
253. Korshunova Y, Maloney RK, Lakey N, Citek RW, Bacher B, Budiman A, et al. Massively parallel bisulphite pyrosequencing reveals the molecular complexity of breast cancer-associated cytosine-methylation patterns obtained from tissue and serum DNA. Genome Res. 2008 Jan;18(1):19-29.
254. Fan JB, Gunderson KL, Bibikova M, Yeakley JM, Chen J, Wickham Garcia E, et al. Illumina universal bead arrays. Methods Enzymol. 2006;410:57-73.
255. Bibikova M, Lin Z, Zhou L, Chudin E, Garcia EW, Wu B, et al. High-throughput DNA methylation profiling using universal bead arrays. Genome Res. 2006 Mar;16(3):383-93.
256. Santos F, Hendrich B, Reik W, Dean W. Dynamic reprogramming of DNA methylation in the early mouse embryo. Dev Biol. 2002 Jan 1;241(1):172-82.
257. Gaudet F, Rideout WM, 3rd, Meissner A, Dausman J, Leonhardt H, Jaenisch R. Dnmt1 expression in pre- and postimplantation embryogenesis and the maintenance of IAP silencing. Mol Cell Biol. 2004 Feb;24(4):1640-8.
258. Suzuki MM, Bird A. DNA methylation landscapes: provocative insights from epigenomics. Nat Rev Genet. 2008 Jun;9(6):465-76.
259. Bird AP. CpG-rich islands and the function of DNA methylation. Nature. 1986 May 15-21;321(6067):209-13.
260. Bartolomei MS, Tilghman SM. Genomic imprinting in mammals. Annu Rev Genet. 1997;31:493-525.
261. Strathdee G, Sim A, Soutar R, Holyoake TL, Brown R. HOXA5 is targeted by cell-type-specific CpG island methylation in normal cells and during the development of acute myeloid leukaemia. Carcinogenesis. 2007 Feb;28(2):299-309.
262. Xu GL, Bestor TH, Bourc'his D, Hsieh CL, Tommerup N, Bugge M, et al. Chromosome instability and immunodeficiency syndrome caused by mutations in a DNA methyltransferase gene. Nature. 1999 Nov 11;402(6758):187-91.
263. Robertson KD. DNA methylation and human disease. Nat Rev Genet. 2005 Aug;6(8):597-610.
264. Feinberg AP. Phenotypic plasticity and the epigenetics of human disease. Nature. 2007 May 24;447(7143):433-40.
265. Fraga MF, Ballestar E, Paz MF, Ropero S, Setien F, Ballestar ML, et al. Epigenetic differences arise during the lifetime of monozygotic twins. Proc Natl Acad Sci U S A. 2005 Jul 26;102(30):10604-9.
266. Issa JP, Ottaviano YL, Celano P, Hamilton SR, Davidson NE, Baylin SB. Methylation of the oestrogen receptor CpG island links ageing and neoplasia in human colon. Nat Genet. 1994 Aug;7(4):536-40.
267. Feinberg AP, Vogelstein B. Hypomethylation distinguishes genes of some human cancers from their normal counterparts. Nature. 1983 Jan 6;301(5895):89-92.
268. Yoder JA, Walsh CP, Bestor TH. Cytosine methylation and the ecology of intragenomic parasites. Trends Genet. 1997 Aug;13(8):335-40.
269. Cui H, Onyango P, Brandenburg S, Wu Y, Hsieh CL, Feinberg AP. Loss of imprinting in colorectal cancer linked to hypomethylation of H19 and IGF2. Cancer Res. 2002 Nov 15;62(22):6442-6.

270. Jones PA, Baylin SB. The fundamental role of epigenetic events in cancer. Nat Rev Genet. 2002 Jun;3(6):415-28.
271. Baylin SB, Ohm JE. Epigenetic gene silencing in cancer - a mechanism for early oncogenic pathway addiction? Nat Rev Cancer. 2006 Feb;6(2):107-16.
272. Shen L, Kondo Y, Rosner GL, Xiao L, Hernandez NS, Vilaythong J, et al. MGMT promoter methylation and field defect in sporadic colorectal cancer. J Natl Cancer Inst. 2005 Sep 21;97(18):1330-8.
273. Yan PS, Venkataramu C, Ibrahim A, Liu JC, Shen RZ, Diaz NM, et al. Mapping geographic zones of cancer risk with epigenetic biomarkers in normal breast tissue. Clin Cancer Res. 2006 Nov 15;12(22):6626-36.
274. Jones PA, Baylin SB. The epigenomics of cancer. Cell. 2007 Feb 23;128(4):683-92.
275. Bird A. Perceptions of epigenetics. Nature. 2007 May 24;447(7143):396-8.
276. Issa JP, Ahuja N, Toyota M, Bronner MP, Brentnall TA. Accelerated age-related CpG island methylation in ulcerative colitis. Cancer Res. 2001 May 1;61(9):3573-7.
277. Nguyen C, Liang G, Nguyen TT, Tsao-Wei D, Groshen S, Lubbert M, et al. Susceptibility of nonpromoter CpG islands to de novo methylation in normal and neoplastic cells. J Natl Cancer Inst. 2001 Oct 3;93(19):1465-72.
278. Frigola J, Song J, Stirzaker C, Hinshelwood RA, Peinado MA, Clark SJ. Epigenetic remodeling in colorectal cancer results in coordinate gene suppression across an entire chromosome band. Nat Genet. 2006 May;38(5):540-9.
279. Ohm JE, McGarvey KM, Yu X, Cheng L, Schuebel KE, Cope L, et al. A stem cell-like chromatin pattern may predispose tumor suppressor genes to DNA hypermethylation and heritable silencing. Nat Genet. 2007 Feb;39(2):237-42.
280. Widschwendter M, Fiegl H, Egle D, Mueller-Holzner E, Spizzo G, Marth C, et al. Epigenetic stem cell signature in cancer. Nat Genet. 2007 Feb;39(2):157-8.
281. Szyf M, Pakneshan P, Rabbani SA. DNA methylation and breast cancer. Biochem Pharmacol. 2004 Sep 15;68(6):1187-97.
282. Yang X, Yan L, Davidson NE. DNA methylation in breast cancer. Endocr Relat Cancer. 2001 Jun;8(2):115-27.
283. Brooks J, Cairns P, Zeleniuch-Jacquotte A. Promoter methylation and the detection of breast cancer. Cancer Causes Control. 2009 Nov;20(9):1539-50.
284. Agrawal A, Murphy RF, Agrawal DK. DNA methylation in breast and colorectal cancers. Mod Pathol. 2007 Jul;20(7):711-21.
285. Evron E, Dooley WC, Umbricht CB, Rosenthal D, Sacchi N, Gabrielson E, et al. Detection of breast cancer cells in ductal lavage fluid by methylation-specific PCR. Lancet. 2001 Apr 28;357(9265):1335-6.
286. Fackler MJ, McVeigh M, Evron E, Garrett E, Mehrotra J, Polyak K, et al. DNA methylation of RASSF1A, HIN-1, RAR-beta, Cyclin D2 and Twist in in situ and invasive lobular breast carcinoma. Int J Cancer. 2003 Dec 20;107(6):970-5.
287. Fackler MJ, McVeigh M, Mehrotra J, Blum MA, Lange J, Lapides A, et al. Quantitative multiplex methylation-specific PCR assay for the detection of promoter hypermethylation in multiple genes in breast cancer. Cancer Res. 2004 Jul 1;64(13):4442-52.
288. Fackler MJ, Rivers A, Teo WW, Mangat A, Taylor E, Zhang Z, et al. Hypermethylated genes as biomarkers of cancer in women with pathologic nipple discharge. Clin Cancer Res. 2009 Jun 1;15(11):3802-11.
289. Mirza S, Sharma G, Prasad CP, Parshad R, Srivastava A, Gupta SD, et al. Promoter hypermethylation of TMS1, BRCA1, ERalpha and PRB in serum and

tumor DNA of invasive ductal breast carcinoma patients. Life Sci. 2007 Jul 4;81(4):280-7.
290. Sharma G, Mirza S, Prasad CP, Srivastava A, Gupta SD, Ralhan R. Promoter hypermethylation of p16INK4A, p14ARF, CyclinD2 and Slit2 in serum and tumor DNA from breast cancer patients. Life Sci. 2007 Apr 24;80(20):1873-81.
291. Dulaimi E, Hillinck J, Ibanez de Caceres I, Al-Saleem T, Cairns P. Tumor suppressor gene promoter hypermethylation in serum of breast cancer patients. Clin Cancer Res. 2004 Sep 15;10(18 Pt 1):6189-93.
292. Jeronimo C, Monteiro P, Henrique R, Dinis-Ribeiro M, Costa I, Costa VL, et al. Quantitative hypermethylation of a small panel of genes augments the diagnostic accuracy in fine-needle aspirate washings of breast lesions. Breast Cancer Res Treat. 2008 May;109(1):27-34.
293. Muller HM, Fiegl H, Widschwendter A, Widschwendter M. Prognostic DNA methylation marker in serum of cancer patients. Ann N Y Acad Sci. 2004 Jun;1022:44-9.
294. Elsheikh SE, Green AR, Rakha EA, Powe DG, Ahmed RA, Collins HM, et al. Global histone modifications in breast cancer correlate with tumor phenotypes, prognostic factors, and patient outcome. Cancer Res. 2009 May 1;69(9):3802-9.
295. Kioulafa M, Kaklamanis L, Stathopoulos E, Mavroudis D, Georgoulias V, Lianidou ES. Kallikrein 10 (KLK10) methylation as a novel prognostic biomarker in early breast cancer. Ann Oncol. 2009 Jun;20(6):1020-5.
296. Kioulafa M, Balkouranidou I, Sotiropoulou G, Kaklamanis L, Mavroudis D, Georgoulias V, et al. Methylation of cystatin M promoter is associated with unfavorable prognosis in operable breast cancer. Int J Cancer. 2009 Jun 23.
297. Kioulafa M, Kaklamanis L, Mavroudis D, Georgoulias V, Lianidou ES. Prognostic significance of RASSF1A promoter methylation in operable breast cancer. Clin Biochem. 2009 Jul;42(10-11):970-5.
298. Hartmann O, Spyratos F, Harbeck N, Dietrich D, Fassbender A, Schmitt M, et al. DNA methylation markers predict outcome in node-positive, estrogen receptor-positive breast cancer with adjuvant anthracycline-based chemotherapy. Clin Cancer Res. 2009 Jan 1;15(1):315-23.
299. Jing F, Jun L, Yong Z, Wang Y, Fei X, Zhang J, et al. Multigene methylation in serum of sporadic Chinese female breast cancer patients as a prognostic biomarker. Oncology. 2008;75(1-2):60-6.
300. Veeck J, Geisler C, Noetzel E, Alkaya S, Hartmann A, Knuchel R, et al. Epigenetic inactivation of the secreted frizzled-related protein-5 (SFRP5) gene in human breast cancer is associated with unfavorable prognosis. Carcinogenesis. 2008 May;29(5):991-8.
301. Nimmrich I, Sieuwerts AM, Meijer-van Gelder ME, Schwope I, Bolt-de Vries J, Harbeck N, et al. DNA hypermethylation of PITX2 is a marker of poor prognosis in untreated lymph node-negative hormone receptor-positive breast cancer patients. Breast Cancer Res Treat. 2008 Oct;111(3):429-37.
302. Maier S, Nimmrich I, Koenig T, Eppenberger-Castori S, Bohlmann I, Paradiso A, et al. DNA-methylation of the homeodomain transcription factor PITX2 reliably predicts risk of distant disease recurrence in tamoxifen-treated, node-negative breast cancer patients--Technical and clinical validation in a multi-centre setting in collaboration with the European Organisation for Research and Treatment of Cancer (EORTC) PathoBiology group. Eur J Cancer. 2007 Jul;43(11):1679-86.

303. Noetzel E, Veeck J, Niederacher D, Galm O, Horn F, Hartmann A, et al. Promoter methylation-associated loss of ID4 expression is a marker of tumour recurrence in human breast cancer. BMC Cancer. 2008;8:154.
304. Fiegl H, Millinger S, Mueller-Holzner E, Marth C, Ensinger C, Berger A, et al. Circulating tumor-specific DNA: a marker for monitoring efficacy of adjuvant therapy in cancer patients. Cancer Res. 2005 Feb 15;65(4):1141-5.
305. Fiegl H, Jones A, Hauser-Kronberger C, Hutarew G, Reitsamer R, Jones RL, et al. Methylated NEUROD1 promoter is a marker for chemosensitivity in breast cancer. Clin Cancer Res. 2008 Jun 1;14(11):3494-502.
306. Martens JW, Nimmrich I, Koenig T, Look MP, Harbeck N, Model F, et al. Association of DNA methylation of phosphoserine aminotransferase with response to endocrine therapy in patients with recurrent breast cancer. Cancer Res. 2005 May 15;65(10):4101-17.
307. Widschwendter M, Siegmund KD, Muller HM, Fiegl H, Marth C, Muller-Holzner E, et al. Association of breast cancer DNA methylation profiles with hormone receptor status and response to tamoxifen. Cancer Res. 2004 Jun 1;64(11):3807-13.
308. Gulliford MC, Bell J, Bourne HM, Petruckevitch A. The reliability of Cancer Registry records. Br J Cancer. 1993 Apr;67(4):819-21.
309. Lapham R, Waugh NR. An audit of the quality of cancer registration data. Br J Cancer. 1992 Sep;66(3):552-4.
310. Brewster DH, Stockton D, Harvey J, Mackay M. Reliability of cancer registration data in Scotland, 1997. Eur J Cancer. 2002 Feb;38(3):414-7.
311. Abraham L, Geller BM, Yankaskas BC, Bowles EJ, Karliner LS, Taplin SH, et al. Accuracy of self-reported breast cancer among women undergoing mammography. Breast Cancer Res Treat. 2009 Mar 20.
312. Parikh-Patel A, Allen M, Wright WE. Validation of self-reported cancers in the California Teachers Study. Am J Epidemiol. 2003 Mar 15;157(6):539-45.
313. MacDonald N, Sibley K, Rosenthal A, Menon U, Jayarajah A, Oram D, et al. A comparison of national cancer registry and direct follow-up in the ascertainment of ovarian cancer. Br J Cancer. 1999 Aug;80(11):1826-7.
314. Paganini-Hill A, Chao A. Accuracy of recall of hip fracture, heart attack, and cancer: a comparison of postal survey data and medical records. Am J Epidemiol. 1993 Jul 15;138(2):101-6.
315. Schrijvers CT, Stronks K, van de Mheen DH, Coebergh JW, Mackenbach JP. Validation of cancer prevalence data from a postal survey by comparison with cancer registry records. Am J Epidemiol. 1994 Feb 15;139(4):408-14.
316. Bergmann MM, Calle EE, Mervis CA, Miracle-McMahill HL, Thun MJ, Heath CW. Validity of self-reported cancers in a prospective cohort study in comparison with data from state cancer registries. Am J Epidemiol. 1998 Mar 15;147(6):556-62.
317. Desai MM, Bruce ML, Desai RA, Druss BG. Validity of self-reported cancer history: a comparison of health interview data and cancer registry records. Am J Epidemiol. 2001 Feb 1;153(3):299-306.
318. Manjer J, Merlo J, Berglund G. Validity of self-reported information on cancer: determinants of under- and over-reporting. Eur J Epidemiol. 2004;19(3):239-47.
319. Dominguez FJ, Lawrence C, Halpern EF, Drohan B, Grinstein G, Black DM, et al. Accuracy of self-reported personal history of cancer in an outpatient breast center. J Genet Couns. 2007 Jun;16(3):341-5.

320. Brewster DH, Stockton DL. Ascertainment of breast cancer by the Scottish Cancer Registry: an assessment based on comparison with five independent breast cancer trials databases. Breast. 2008 Feb;17(1):104-6.
321. Schootman M JD, West MM, Aft R. Schootman et al,. Self-reported by elderly breast cancer patients was an acceptable altenative to surveillance, epidemiology, and end results (SEER) abstract data. . J Clin Epidemiol 2005;58(12):1316-9.
322. Phillips KA, Milne RL, Buys S, Friedlander ML, Ward JH, McCredie MR, et al. Agreement between self-reported breast cancer treatment and medical records in a population-based Breast Cancer Family Registry. J Clin Oncol. 2005 Jul 20;23(21):4679-86.
323. Downing A, Prakash K, Gilthorpe MS, Mikeljevic JS, Forman D. Socioeconomic background in relation to stage at diagnosis, treatment and survival in women with breast cancer. Br J Cancer. 2007 Mar 12;96(5):836-40.
324. Dawson SJ, Duffy SW, Blows FM, Driver KE, Provenzano E, LeQuesne J, et al. Molecular characteristics of screen-detected vs symptomatic breast cancers and their impact on survival. Br J Cancer. 2009 Oct 20;101(8):1338-44.
325. Blamey RW, Hornmark-Stenstam B, Ball G, Blichert-Toft M, Cataliotti L, Fourquet A, et al. ONCOPOOL - a European database for 16,944 cases of breast cancer. Eur J Cancer. 2010 Jan;46(1):56-71.
326. Pascoe SW, Neal RD, Heywood PL, Allgar VL, Miles JN, Stefoski-Mikeljevic J. Identifying patients with a cancer diagnosis using general practice medical records and Cancer Registry data. Fam Pract. 2008 Aug;25(4):215-20.
327. Kaye JA, Derby LE, del Mar Melero-Montes M, Quinn M, Jick H. The incidence of breast cancer in the General Practice Research Database compared with national cancer registration data. Br J Cancer. 2000 Dec;83(11):1556-8.
328. Jensen AR, Overgaard J, Storm HH. Validity of breast cancer in the Danish Cancer Registry. A study based on clinical records from one county in Denmark. Eur J Cancer Prev. 2002 Aug;11(4):359-64.
329. Pearson ML, Ganz PA, McGuigan K, Malin JR, Adams J, Kahn KL. The case identification challenge in measuring quality of cancer care. J Clin Oncol. 2002 Nov 1;20(21):4353-60.
330. Gathani T, Bull D, Green J, Reeves G, Beral V. Breast cancer histological classification: agreement between the Office for National Statistics and the National Health Service Breast Screening Programme. Breast Cancer Res. 2005;7(6):R1090-6.
331. Stotter A, Bright N, Silcocks PB, Botha JL. Effect of improved data collection on breast cancer incidence and survival: reconciliation of a registry with a clinical database. Bmj. 2000 Jul 22;321(7255):214.
332. Villard-Mackintosh L, Coleman MP, Vessey MP. The completeness of cancer registration in England: an assessment from the Oxford-FPA contraceptive study. Br J Cancer. 1988 Oct;58(4):507-11.
333. Hunt K, Coleman MP. The completeness of cancer registration in follow-up studies--a cautionary note. Br J Cancer. 1987 Sep;56(3):357-9.
334. Key TJ, Verkasalo PK, Banks E. Epidemiology of breast cancer. Lancet Oncol. 2001 Mar;2(3):133-40.
335. Menon U, Gentry-Maharaj A, Hallett R, Ryan A, Burnell M, Sharma A, et al. Sensitivity and specificity of multimodal and ultrasound screening for ovarian cancer, and stage distribution of detected cancers: results of the prevalence screen of the UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS). Lancet Oncol. 2009 Mar 10.

336. Vermeulen A, Verdonck L, Kaufman JM. A critical evaluation of simple methods for the estimation of free testosterone in serum. J Clin Endocrinol Metab. 1999 Oct;84(10):3666-72.
337. Sievernich A, Wildt L, Lichtenberg-Frate H. In vitro bioactivity of 17alpha-estradiol. J Steroid Biochem Mol Biol. 2004 Dec;92(5):455-63.
338. Hasenbrink G, Sievernich A, Wildt L, Ludwig J, Lichtenberg-Frate H. Estrogenic effects of natural and synthetic compounds including tibolone assessed in Saccharomyces cerevisiae expressing the human estrogen alpha and beta receptors. Faseb J. 2006 Jul;20(9):1552-4.
339. Miller S, Kennedy D, Thomson J, Han F, Smith R, Ing N, et al. A rapid and sensitive reporter gene that uses green fluorescent protein expression to detect chemicals with estrogenic activity. Toxicol Sci. 2000 May;55(1):69-77.
340. Lippman M, Monaco ME, Bolan G. Effects of estrone, estradiol, and estriol on hormone-responsive human breast cancer in long-term tissue culture. Cancer Res. 1977 Jun;37(6):1901-7.
341. Aronica SM, Katzenellenbogen BS. Stimulation of estrogen receptor-mediated transcription and alteration in the phosphorylation state of the rat uterine estrogen receptor by estrogen, cyclic adenosine monophosphate, and insulin-like growth factor-I. Mol Endocrinol. 1993 Jun;7(6):743-52.
342. Key TJ, Appleby PN, Reeves GK, Roddam AW. Insulin-like growth factor 1 (IGF1), IGF binding protein 3 (IGFBP3), and breast cancer risk: pooled individual data analysis of 17 prospective studies. Lancet Oncol. 2010 Jun;11(6):530-42.
343. Harris TG, Strickler HD, Yu H, Pollak MN, Monrad ES, Travin MI, et al. Specimen processing time and measurement of total insulin-like growth factor-I (IGF-I), free IGF-I, and IGF binding protein-3 (IGFBP-3). Growth Horm IGF Res. 2006 Apr;16(2):86-92.
344. O'Malley BW. A life-long search for the molecular pathways of steroid hormone action. Mol Endocrinol. 2005 Jun;19(6):1402-11.
345. Fox EM, Davis RJ, Shupnik MA. ERbeta in breast cancer--onlooker, passive player, or active protector? Steroids. 2008 Oct;73(11):1039-51.
346. Judd HL, Shamonki IM, Frumar AM, Lagasse LD. Origin of serum estradiol in postmenopausal women. Obstet Gynecol. 1982 Jun;59(6):680-6.
347. Garland CF, Friedlander NJ, Barrett-Connor E, Khaw KT. Sex hormones and postmenopausal breast cancer: a prospective study in an adult community. Am J Epidemiol. 1992 Jun 1;135(11):1220-30.
348. Rinaldi S, Geay A, Dechaud H, Biessy C, Zeleniuch-Jacquotte A, Akhmedkhanov A, et al. Validity of free testosterone and free estradiol determinations in serum samples from postmenopausal women by theoretical calculations. Cancer Epidemiol Biomarkers Prev. 2002 Oct;11(10 Pt 1):1065-71.
349. Anderson JN, Peck EJ, Jr., Clark JH. Estrogen-induced uterine responses and growth: relationship to receptor estrogen binding by uterine nuclei. Endocrinology. 1975 Jan;96(1):160-7.
350. Maggiolini M, Vivacqua A, Fasanella G, Recchia AG, Sisci D, Pezzi V, et al. The G protein-coupled receptor GPR30 mediates c-fos up-regulation by 17beta-estradiol and phytoestrogens in breast cancer cells. J Biol Chem. 2004 Jun 25;279(26):27008-16.
351. Yager JD. Endogenous estrogens as carcinogens through metabolic activation. J Natl Cancer Inst Monogr. 2000(27):67-73.
352. Cox DG, Blanche H, Pearce CL, Calle EE, Colditz GA, Pike MC, et al. A comprehensive analysis of the androgen receptor gene and risk of breast cancer: results from the National Cancer Institute Breast and Prostate Cancer Cohort Consortium (BPC3). Breast Cancer Res. 2006;8(5):R54.

353. Maggiolini M, Donze O, Jeannin E, Ando S, Picard D. Adrenal androgens stimulate the proliferation of breast cancer cells as direct activators of estrogen receptor alpha. Cancer Res. 1999 Oct 1;59(19):4864-9.
354. Tamimi RM, Hankinson SE, Chen WY, Rosner B, Colditz GA. Combined estrogen and testosterone use and risk of breast cancer in postmenopausal women. Arch Intern Med. 2006 Jul 24;166(14):1483-9.
355. Kenemans P, Bundred NJ, Foidart JM, Kubista E, von Schoultz B, Sismondi P, et al. Safety and efficacy of tibolone in breast-cancer patients with vasomotor symptoms: a double-blind, randomised, non-inferiority trial. Lancet Oncol. 2009 Feb;10(2):135-46.
356. Lange CA. Challenges to defining a role for progesterone in breast cancer. Steroids. 2008 Oct;73(9-10):914-21.
357. Menon U, Burnell M, Sharma A, Gentry-Maharaj A, Fraser L, Ryan A, et al. Decline in use of hormone therapy among postmenopausal women in the United Kingdom. Menopause-the Journal of the North American Menopause Society. 2007 May-Jun;14(3):462-7.
358. Yu H, Shu XO, Shi R, Dai Q, Jin F, Gao YT, et al. Plasma sex steroid hormones and breast cancer risk in Chinese women. Int J Cancer. 2003 May 20;105(1):92-7.
359. Boyd NF, Stone J, Martin LJ, Jong R, Fishell E, Yaffe M, et al. The association of breast mitogens with mammographic densities. Br J Cancer. 2002 Oct 7;87(8):876-82.
360. Johansson H, Gandini S, Bonanni B, Mariette F, Guerrieri-Gonzaga A, Serrano D, et al. Relationships between circulating hormone levels, mammographic percent density and breast cancer risk factors in postmenopausal women. Breast Cancer Research and Treatment. 2008 Mar;108(1):57-67.
361. Wang DY, Goodwin PR, Bulbrook RD, Hayward JL. Plasma FSH and LH in post-menopausal women with breast cancer. Eur J Cancer. 1976 Apr;12(4):305-11.
362. Armstrong PCKLaDT. Interaction of steroids and gonsdotropins in the control of steroidogenesis in the ovarian follicle. Ann Rev Physiol. 1980;42:71-82.
363. Barnes RB, Rosenfield RL, Namnoum A, Layman LC. Effect of follicle-stimulating hormone on ovarian androgen production in a woman with isolated follicle-stimulating hormone deficiency. N Engl J Med. 2000 Oct 19;343(16):1197-8.
364. (http://www.ncbi.nlm.nih.gov/geo).
365. Mani AM, Fenwick MA, Cheng ZR, Sharma MK, Singh D, Wathes DC. IGF1 induces up-regulation of steroidogenic and apoptotic regulatory genes via activation of phosphatidylinositol-dependent kinase/AKT in bovine granulosa cells. Reproduction. 2010 Jan;139(1):139-51.
366. Luo WX, Wiltbank MC. Distinct regulation by steroids of messenger RNAs for FSHR and CYP19A1 in bovine granulosa cells. Biology of Reproduction. 2006 Aug;75(2):217-25.
367. Chodankar R, Kwang S, Sangiorgi F, Hong H, Yen HY, Deng C, et al. Cell-nonautonomous induction of ovarian and uterine serous cystadenomas in mice lacking a functional Brca1 in ovarian granulosa cells. Curr Biol. 2005 Mar 29;15(6):561-5.
368. GH H. Tumor heterogeneity. Cancer Research. 1984;44:2259-65.
369. Khalique L, Ayhan A, Weale ME, Jacobs IJ, Ramus SJ, Gayther SA. Genetic intra-tumour heterogeneity in epithelial ovarian cancer and its implications for molecular diagnosis of tumours. J Pathol. 2007 Feb;211(3):286-95.

370. Losi L, Baisse B, Bouzourene H, Benhattar J. Evolution of intratumoral genetic heterogeneity during colorectal cancer progression. Carcinogenesis. 2005 May;26(5):916-22.
371. Lyng MB, Laenkholm AV, Pallisgaard N, Vach W, Knoop A, Bak M, et al. Intratumor genetic heterogeneity of breast carcinomas as determined by fine needle aspiration and TaqMan low density array. Cell Oncol. 2007;29(5):361-72.
372. Ruiz-Cerda JL, Hernandez M, Sempere A, O'Connor JE, Kimler BF, Jimenez-Cruz F. Intratumoral heterogeneity of DNA content in renal cell carcinoma and its prognostic significance. Cancer. 1999 Aug 15;86(4):664-71.
373. van der Poel HG, Oosterhof GO, Schaafsma HE, Debruyne FM, Schalken JA. Intratumoral nuclear morphologic heterogeneity in prostate cancer. Urology. 1997 Apr;49(4):652-7.
374. Lyng H, Beigi M, Svendsrud DH, Brustugun OT, Stokke T, Kristensen GB, et al. Intratumor chromosomal heterogeneity in advanced carcinomas of the uterine cervix. Int J Cancer. 2004 Sep 1;111(3):358-66.
375. Aubele M, Mattis A, Zitzelsberger H, Walch A, Kremer M, Hutzler P, et al. Intratumoral heterogeneity in breast carcinoma revealed by laser-microdissection and comparative genomic hybridization. Cancer Genet Cytogenet. 1999 Apr 15;110(2):94-102.
376. Zhu G, Reynolds L, Crnogorac-Jurcevic T, Gillett CE, Dublin EA, Marshall JF, et al. Combination of microdissection and microarray analysis to identify gene expression changes between differentially located tumour cells in breast cancer. Oncogene. 2003 Jun 12;22(24):3742-8.
377. Nakamura T, Kuwai T, Kitadai Y, Sasaki T, Fan D, Coombes KR, et al. Zonal heterogeneity for gene expression in human pancreatic carcinoma. Cancer Res. 2007 Aug 15;67(16):7597-604.
378. Sigalotti L, Coral S, Nardi G, Spessotto A, Cortini E, Cattarossi I, et al. Promoter methylation controls the expression of MAGE2, 3 and 4 genes in human cutaneous melanoma. J Immunother. 2002 Jan-Feb;25(1):16-26.
379. Kuniyasu H, Troncoso P, Johnston D, Bucana CD, Tahara E, Fidler IJ, et al. Relative expression of type IV collagenase, E-cadherin, and vascular endothelial growth factor/vascular permeability factor in prostatectomy specimens distinguishes organ-confined from pathologically advanced prostate cancers. Clin Cancer Res. 2000 Jun;6(6):2295-308.
380. Jarque F, Lluch A, Vera FJ, Pascual A, Vizcarra E, Alberola V, et al. Intratumoral variation of estrogen and progesterone receptors in breast cancer: relationship with histopathological characteristics of the tumor. Oncology. 1990;47(1):9-13.
381. Davis BW, Zava DT, Locher GW, Goldhirsch A, Hartmann WH. Receptor heterogeneity of human breast cancer as measured by multiple intratumoral assays of estrogen and progesterone receptor. Eur J Cancer Clin Oncol. 1984 Mar;20(3):375-82.
382. Van Notten JP AJ, Caryle SS, Goodchild NL, Thornton IG, Brisgen ML, Coy P, Fletcher C. Estrogen receptor distribution in the peirpheral, intermediate and central regions of breast cancer. Eur J Cancer Clin Oncol. 1988;24(1885-1889).
383. Widschwendter M, Berger J, Daxenbichler G, Muller-Holzner E, Widschwendter A, Mayr A, et al. Loss of retinoic acid receptor beta expression in breast cancer and morphologically normal adjacent tissue but not in the normal breast tissue distant from the cancer. Cancer Res. 1997 Oct 1;57(19):4158-61.
384. Widschwendter M, Berger J, Hermann M, Muller HM, Amberger A, Zeschnigk M, et al. Methylation and silencing of the retinoic acid receptor-beta2 gene in breast cancer. J Natl Cancer Inst. 2000 May 17;92(10):826-32.

385. Bistulfi G, Pozzi S, Ren M, Rossetti S, Sacchi N. A repressive epigenetic domino effect confers susceptibility to breast epithelial cell transformation: implications for predicting breast cancer risk. Cancer Res. 2006 Nov 1;66(21):10308-14.
386. Lewis JT, Ketterling RP, Halling KC, Reynolds C, Jenkins RB, Visscher DW. Analysis of intratumoral heterogeneity and amplification status in breast carcinomas with equivocal (2+) HER-2 immunostaining. Am J Clin Pathol. 2005 Aug;124(2):273-81.
387. Laird PW. The power and the promise of DNA methylation markers. Nat Rev Cancer. 2003 Apr;3(4):253-66.
388. Schlesinger Y, Straussman R, Keshet I, Farkash S, Hecht M, Zimmerman J, et al. Polycomb-mediated methylation on Lys27 of histone H3 pre-marks genes for de novo methylation in cancer. Nat Genet. 2007 Feb;39(2):232-6.
389. Henderson GS, van Diest PJ, Burger H, Russo J, Raman V. Expression pattern of a homeotic gene, HOXA5, in normal breast and in breast tumors. Cell Oncol. 2006;28(5-6):305-13.
390. Chen H, Sukumar S. HOX genes: emerging stars in cancer. Cancer Biol Ther. 2003 Sep-Oct;2(5):524-5.
391. Leu YW, Yan PS, Fan M, Jin VX, Liu JC, Curran EM, et al. Loss of estrogen receptor signaling triggers epigenetic silencing of downstream targets in breast cancer. Cancer Res. 2004 Nov 15;64(22):8184-92.
392. Bieche I, Franc B, Vidaud D, Vidaud M, Lidereau R. Analyses of MYC, ERBB2, and CCND1 genes in benign and malignant thyroid follicular cell tumors by real-time polymerase chain reaction. Thyroid. 2001 Feb;11(2):147-52.
393. Harbeck N, Nimmrich I, Hartmann A, Ross JS, Cufer T, Grutzmann R, et al. Multicenter study using paraffin-embedded tumor tissue testing PITX2 DNA methylation as a marker for outcome prediction in tamoxifen-treated, node-negative breast cancer patients. J Clin Oncol. 2008 Nov 1;26(31):5036-42.
394. Sigalotti L, Fratta E, Coral S, Tanzarella S, Danielli R, Colizzi F, et al. Intratumor heterogeneity of cancer/testis antigens expression in human cutaneous melanoma is methylation-regulated and functionally reverted by 5-aza-2'-deoxycytidine. Cancer Res. 2004 Dec 15;64(24):9167-71.
395. Rastetter M, Schagdarsurengin U, Lahtz C, Fiedler E, Marsch W, Dammann R, et al. Frequent intra-tumoural heterogeneity of promoter hypermethylation in malignant melanoma. Histol Histopathol. 2007 Sep;22(9):1005-15.
396. Woloszynska-Read A, Mhawech-Fauceglia P, Yu J, Odunsi K, Karpf AR. Intertumor and intratumor NY-ESO-1 expression heterogeneity is associated with promoter-specific and global DNA methylation status in ovarian cancer. Clin Cancer Res. 2008 Jun 1;14(11):3283-90.
397. Ebert MP, Mooney SH, Tonnes-Priddy L, Lograsso J, Hoffmann J, Chen J, et al. Hypermethylation of the TPEF/HPP1 gene in primary and metastatic colorectal cancers. Neoplasia. 2005 Aug;7(8):771-8.
398. Liang G, Robertson KD, Talmadge C, Sumegi J, Jones PA. The gene for a novel transmembrane protein containing epidermal growth factor and follistatin domains is frequently hypermethylated in human tumor cells. Cancer Res. 2000 Sep 1;60(17):4907-12.
399. Shibata DM, Sato F, Mori Y, Perry K, Yin J, Wang S, et al. Hypermethylation of HPP1 is associated with hMLH1 hypermethylation in gastric adenocarcinomas. Cancer Res. 2002 Oct 15;62(20):5637-40.
400. Goo YA, Goodlett DR, Pascal LE, Worthington KD, Vessella RL, True LD, et al. Stromal mesenchyme cell genes of the human prostate and bladder. BMC Urol. 2005;5:17.

401. Lind GE, Kleivi K, Meling GI, Teixeira MR, Thiis-Evensen E, Rognum TO, et al. ADAMTS1, CRABP1, and NR3C1 identified as epigenetically deregulated genes in colorectal tumorigenesis. Cell Oncol. 2006;28(5-6):259-72.
402. Wu Q, Lothe RA, Ahlquist T, Silins I, Trope CG, Micci F, et al. DNA methylation profiling of ovarian carcinomas and their in vitro models identifies HOXA9, HOXB5, SCGB3A1, and CRABP1 as novel targets. Mol Cancer. 2007;6:45.
403. Piotrowski A, Benetkiewicz M, Menzel U, de Stahl TD, Mantripragada K, Grigelionis G, et al. Microarray-based survey of CpG islands identifies concurrent hyper- and hypomethylation patterns in tissues derived from patients with breast cancer. Genes Chromosomes Cancer. 2006 Jul;45(7):656-67.
404. Pearson JC, Lemons D, McGinnis W. Modulating Hox gene functions during animal body patterning. Nat Rev Genet. 2005 Dec;6(12):893-904.
405. Care A, Silvani A, Meccia E, Mattia G, Stoppacciaro A, Parmiani G, et al. HOXB7 constitutively activates basic fibroblast growth factor in melanomas. Mol Cell Biol. 1996 Sep;16(9):4842-51.
406. Ordway JM, Budiman MA, Korshunova Y, Maloney RK, Bedell JA, Citek RW, et al. Identification of novel high-frequency DNA methylation changes in breast cancer. PLoS One. 2007;2(12):e1314.
407. Novak P, Jensen T, Oshiro MM, Wozniak RJ, Nouzova M, Watts GS, et al. Epigenetic inactivation of the HOXA gene cluster in breast cancer. Cancer Res. 2006 Nov 15;66(22):10664-70.
408. Fiegl H, Windbichler G, Mueller-Holzner E, Goebel G, Lechner M, Jacobs IJ, et al. HOXA11 DNA methylation--a novel prognostic biomarker in ovarian cancer. Int J Cancer. 2008 Aug 1;123(3):725-9.
409. Avraham A, Sandbank J, Yarom N, Shalom A, Karni T, Pappo I, et al. A similar cell-specific pattern of HOXA methylation in normal and in cancer tissues. Epigenetics. 2010 Jan;5(1):41-6.
410. Grimshaw MJ, Cooper L, Papazisis K, Coleman JA, Bohnenkamp HR, Chiapero-Stanke L, et al. Mammosphere culture of metastatic breast cancer cells enriches for tumorigenic breast cancer cells. Breast Cancer Res. 2008;10(3):R52.
411. Hwang-Verslues WW, Kuo WH, Chang PH, Pan CC, Wang HH, Tsai ST, et al. Multiple lineages of human breast cancer stem/progenitor cells identified by profiling with stem cell markers. PLoS One. 2009;4(12):e8377.
412. Dworkin AM, Huang TH, Toland AE. Epigenetic alterations in the breast: Implications for breast cancer detection, prognosis and treatment. Semin Cancer Biol. 2009 Jun;19(3):165-71.
413. Teschendorff AE, Menon U, Gentry-Maharaj A, Ramus SJ, Weisenberger DJ, Shen H, et al. Age-dependent DNA methylation of genes that are suppressed in stem cells is a hallmark of cancer. Genome Res. Mar 10.
414. Wu JM, Fackler MJ, Halushka MK, Molavi DW, Taylor ME, Teo WW, et al. Heterogeneity of breast cancer metastases: comparison of therapeutic target expression and promoter methylation between primary tumors and their multifocal metastases. Clin Cancer Res. 2008 Apr 1;14(7):1938-46.
415. Lin HJ, Zuo T, Lin CH, Kuo CT, Liyanarachchi S, Sun S, et al. Breast cancer-associated fibroblasts confer AKT1-mediated epigenetic silencing of Cystatin M in epithelial cells. Cancer Res. 2008 Dec 15;68(24):10257-66.
416. Braakhuis BJ, Tabor MP, Kummer JA, Leemans CR, Brakenhoff RH. A genetic explanation of Slaughter's concept of field cancerization: evidence and clinical implications. Cancer Res. 2003 Apr 15;63(8):1727-30.
417. Fukino K, Shen L, Matsumoto S, Morrison CD, Mutter GL, Eng C. Combined total genome loss of heterozygosity scan of breast cancer stroma and

epithelium reveals multiplicity of stromal targets. Cancer Res. 2004 Oct 15;64(20):7231-6.
418. Qiu W, Hu M, Sridhar A, Opeskin K, Fox S, Shipitsin M, et al. No evidence of clonal somatic genetic alterations in cancer-associated fibroblasts from human breast and ovarian carcinomas. Nat Genet. 2008 May;40(5):650-5.
419. Yazici H, Terry MB, Cho YH, Senie RT, Liao Y, Andrulis I, et al. Aberrant methylation of RASSF1A in plasma DNA before breast cancer diagnosis in the Breast Cancer Family Registry. Cancer Epidemiol Biomarkers Prev. 2009 Oct;18(10):2723-5.
420. Degen M, Brellier F, Schenk S, Driscoll R, Zaman K, Stupp R, et al. Tenascin-W, a new marker of cancer stroma, is elevated in sera of colon and breast cancer patients. Int J Cancer. 2008 Jun 1;122(11):2454-61.
421. Fata JE, Kong YY, Li J, Sasaki T, Irie-Sasaki J, Moorehead RA, et al. The osteoclast differentiation factor osteoprotegerin-ligand is essential for mammary gland development. Cell. 2000 Sep 29;103(1):41-50.
422. Fernandez-Valdivia R, Mukherjee A, Ying Y, Li J, Paquet M, DeMayo FJ, et al. The RANKL signaling axis is sufficient to elicit ductal side-branching and alveologenesis in the mammary gland of the virgin mouse. Dev Biol. 2009 Apr 1;328(1):127-39.

PUBLICATIONS/PRESENTATIONS

Publications

Related to this thesis

Fourkala EO, Hauser-Kronberger C, Apostolidou S, Burnell M, Jones A, Grall J, Reitsamer R, Fiegl H, Jacobs I, Menon U, Widschwendter M. DNA methylation of polycomb group target genes in cores taken from breast cancer centre and periphery Breast Cancer Res Treat, 2010;120(2):345-55.

Jones A, Lechner M, **Fourkala EO**, Kristeleit R, Widschwendter M. Emerging promise of epigenetics and DNA methylation for the diagnosis and management of women's cancer. Epigenomics, 2010; 2(1):9-38.

Fourkala EO, Zaikin A, Burnell M, Gentry-Maharaj A, Ford J, Gunu R, Soromani C, Hasenbrink G, Jacobs I, Dawney A, Widschwendter M, Lichtenberg-Fraté H and Menon U. Serum steroid receptor bioactivity assays and sex steroid hormones along with gonadotrophins predict breast cancer in postmenopausal women. (Manuscript submitted in Breast Cancer Res Treat).

Gentry-Maharaj A*, **Fourkala EO***, Burnell M, Ryan A, Parmar A, Jacobs I, Menon U. Research studies relying on cancer data should collect data from multiple sources. *equally contributed (Manuscript to be submitted in BMJ)

Fourkala EO, Zaikin A, Burnell M, Gentry-Maharaj A, Ford J, Jacobs I, Dawney A, Lichtenberg-Fraté H, Dubeau, Menon U and Widschwendter M. FSH and testosterone act synegestically in breast carcinogenesis (Manuscript in preparation).

Not related to this thesis

Widschwendter M, Apostolidou S, Jones A, **Fourkala EO**, Arora R, Pearce CL, Frasco MA, Ayhan A, Zikan M, Cibula D, Iyibozkurt CA, Yavuz E, Hauser-Kronberger C, Dubeau L, Menon U, Jacobs IJ. HOXA methylation in normal endometrium from premenopausal women is associated with the presence of ovarian cancer: a proof of principle study. Int J Cancer (2009).

Sharma A, Gentry-Maharaj A, Burnell M, **Fourkala EO**, Campbell S, Nazar A, Seif M, Hallett R, Ryan A, Parmar M, Jacobs I, Menon U for United Kingdom Collaborative Trial of Ovarian Cancer Screening (UKCTOCS). Assessing the malignant potential of ovarian inclusion cysts in postmenopausal women-prospective cohosrt study within UKCTOCS. (Submitted in JCO)

Presentations (Oral and Poster)

January 2010: 2nd Student Day, Institute for Women's Health, UCL, London, UK (poster)

November 2010: 2nd Meeting, Innovations & Progress in Healthcare for Women, London, UK (oral and poster presentation)

July 2009: Pattison Review, UCL (2 poster presentations)

Dec 2008: Institute for Women's' Health Annual Meeting, London, UK (awarded 3rd prize for poster presentation)

Oct 2008: Student Day, Institute for Women's Health, UCL, London, UK (oral)

Dec 2007: Institute for Women's' Health Annual Meeting, London, UK (poster)

April 2007: 1st Meeting, Innovations & Progress in Healthcare for Women, London, UK (poster)

Dec 2006: Institute for Women's' Health Annual Meeting, London, UK (poster)

Appendix VI: Statistically non-significant joint associations of high levels of hormones with breast cancer risk

Supplemental Table VI-1: Joint association of high levels of oestrogens (top quantiles) with risk of breast cancer – all cases.

Hormones		Joint association of oestrogens Oestradiol Oestrone
		OR* (95% CI)
Not adjusted		2.016 (1.017-4.668) p=0.072
		OR (95% CI)**
Adjusted	Oestradiol	
	Oestrone	
	Androstenedione	1.884 (0.863-4.111) p=0.108
	Testosterone	1.638 (0.742-3.608) p=0.217
	DHEAS	2.010 (0.928-4.365) p=0.074
	SHBG	1.888 (0.868-4.109) p=0.106
	Progesterone	1.964 (0.906-4.255) p=0.084
	LH	1.946 (0.898-4.222) p=0.089
	FSH	2.036 (0.940-4.415) p=0.069
	ER-α SB	1.822 (0.813-3.994) p=0.132
	ER-β SB	1.863 (0.815-4.077) p=0.117
	AR SB	1.851 (0.811-4.055) p=0.121

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Supplemental Table VI-2: Joint association of high levels of oestrogens and androgens (top quantiles) with risk of breast cancer – all cases.

Hormones		Joint association of oestrogens and androgens		
		Oestradiol	Oestrone	
		DHEAS	Androstenedione	DHEAS
		OR* (95% CI)	**OR* (95% CI)**	**OR* (95% CI)**
Not adjusted		1.090 (0.499-2.148) p=0.817	1.853 (1.037-4.487) p=0.094	1.274 (0.595-2.848) p=0.545
		OR (95% CI)**	**OR** (95% CI)**	**OR** (95% CI)**
Adjusted	Oestradiol		1.768 (0.828-3.739) p=0.135	1.106 (0.476-2.439) p=0.807
	Oestrone	1.060 (0.495-2.174) p=0.877		
	Androstenedione	0.880 (0.397-1.868) p=0.743		0.925 (0.389-2.083) p=0.854
	Testosterone	0.642 (0.286-1.379) p=0.266	1.369 (0.616-3.015) p=0.434	0.750 (0.312-1.709) p=0.503
	DHEAS		1.857 (0.880-3.561) p=0.100	
	SHBG	1.101 (0.512-2.265) p=0.798	1.719 (0.823-4.282) p=0.144	1.149 (0.507-2.496) p=0.730
	Progesterone	1.034 (0.465-2.207) p=0.932	1.977 (0.913-3.691) p=0.081	1.138 (0.486-2.566) p=0.759
	LH	1.074 (0.501-2.199) p=0.849	1.784 (0.855-3.691) p=0.118	1.274 (0.564-2.764) p=0.547
	FSH	1.106 (0.515-2.227) p=0.789	1.852 (0.890-3.822) p=0.094	1.287 (0.570-2.789) p=0.529
	ER-α SB	1.106 (0.517-2.262) p=0.788	1.721 (0.812-3.591) p=0.149	1.283 (0.568-2.778) p=0.534
	ER-β SB	1.095 (0.511-2.241) p=0.809	1.727 (0.972-3.606) p=0.146	1.402 (0.637-2.988) p=0.388
	AR SB	0.988 (0.500-2.054) p=0.974	1.715 (0.917-3.577) p=0.151	1.251 (0.554-2.710) p=0.577

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Supplemental Table VI-3: Joint association of high levels of oestrogens and progesterone (top quantiles) with risk of breast cancer – all cases.

Hormones		Joint association of oestrogens and progesterone	
		Oestradiol	Oestrone
		Progesterone	
		OR* (95% CI)	**OR* (95% CI)**
Not adjusted		1.125 (0.547-2.280) p=0.745	1.670 (0.803-3.712) p=0.187
		OR (95% CI)**	**OR** (95% CI)**
Adjusted	Oestradiol		1.602 (0.713-3.531) p=0.243
	Oestrone	1.194 (0.581-2.377) p=0.618	
	Androstenedione	0.857 (0.393-1.804) p=0.690	1.310 (0.586-2.876) p=0.502
	Testosterone	0.750 (0.342-1.584) p=0.459	0.973 (0.414-2.206) p=0.947
	DHEAS	1.117 (0.518-2.325) p=0.771	1.693 (0.745-3.803) p=0.201
	SHBG	1.234 (0.586-2.511) p=0.568	1.605 (0.734-3.455) p=0.227
	Progesterone		
	LH	1.083 (0.517-2.181) p=0.826	1.635 (0.750-3.505) p=0.207
	FSH	1.139 (0.542-2.302) p=0.722	1.693 (0.777-3.633) p=0.177
	ER-α	1.129 (0.539-2.273) p=0.740	1.525 (0.686-3.309) p=0.289
	ER-β	1.032 (0.483-2.102) p=0.933	1.535 (0.690-3.334) p=0.281
	AR	1.020 (0.476-2.086) p=0.959	1.509 (0.768-3.277) p=0.301

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Supplemental Table VI-4: Joint association of high levels of oestrogens and gonadotrophins (top quantiles) with risk of breast cancer – all cases.

Hormones		Joint association of oestrogens and gonadotrophins			
		Oestradiol		Oestrone	
		LH	FSH	LH	FSH
		OR* (95% CI)	**OR* (95% CI)**	**OR* (95% CI)**	**OR* (95% CI)**
	Not adjusted	0.711 (0.209-2.140) p=0.565	0.714 (0.199-2.930) p=0.622	1.066 (0.405-2.678) p=0.894	1.468 (0.655-3.912) p=0.394
		OR (95% CI)**	**OR** (95% CI)**	**OR** (95% CI)**	**OR** (95% CI)**
Adjusted	Oestradiol			1.052 (0.389-2.671) p=0.915	1.288 (0.496-3.171) p=0.587
	Oestrone	0.776 (0.210-2.366) p=0.673	0.795 (0.170-2.903) p=0.742		
	Androstenedione	0.646 (0.175-1.945) p=0.464	0.687 (0.148-2.437) p=0.586	0.935 (0.322-2.428) p=0.895	1.201 (0.462-3.006) p=0.686
	Testosterone	0.556 (0.149-1.693) p=0.331	0.679 (0.145-2.430) p=0.576	0.945 (0.343-1.398) p=0.908	1.247 (0.475-3.107) p=0.640
	DHEAS	0.704 (0.192-2.098) p=0.553	0.717 (0.155-2.515) p=0.626	1.056 (0.390-2.633) p=0.910	1.469 (0.590-3.537) p=0.393
	SHBG	0.757 (0.205-2.285) p=0.641	0.672 (0.145-2.374) p=0.562	1.014 (0.373-2.541) p=0.978	1.589 (0.630-3.886) p=0.312
	Progesterone	0.696 (0.191-2.073) p=0.541	0.712 (0.154-2.498) p=0.619	1.046 (0.387-2.608) p=0.925	1.469 (0.589-3.538) p=0.394
	LH		0.871 (0.185-3.150) p=0.843		1.661 (0.658-4.067) p=0.268
	FSH	0.697 (0.190-2.087) p=0.544		1.047 (0.385-2.622) p=0.925	
	ER-α SB	0.714 (0.195-2.127) p=0.570	0.718 (0.155-2.522) p=0.628	1.024 (0.377-2.557) p=0.961	1.455 (0.583-3.510) p=0.406
	ER-β SB	0.708 (0.194-2.108) p=0.561	0.715 (0.155-2.508) p=0.624	1.065 (0.394-2.652) p=0.895	1.472 (0.590-3.544) p=0.391
	AR SB	0.705 (0.193-2.099) p=0.555	0.724 (0.156-2.542) p=0.637	1.081 (0.399-2.695) p=0.871	1.490 (0.598-3.590) p=0.376

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Supplemental Table VI-5: Joint association of high levels of oestrogens and gonadotrophins (top quantiles) with risk of breast cancer – all cases.

Hormones		Joint association of oestrogens and SHBG	
		Oestradiol	Oestrone
		SHBG	
		OR* (95% CI)	**OR* (95% CI)**
Not adjusted		0.417 (0.077-1.172) p=0.175	0.468 (0.1404-0.910) p=0.180
		OR (95% CI)**	**OR** (95% CI)**
Adjusted	Oestradiol		0.498 (0.084-1.398) p=0.222
	Oestrone	0.424 (0.096-1.329) p=0.184	
	Androstenedione	0.388 (0.088-1.230) p=0.146	0.537 (0.150-1.519) p=0.279
	Testosterone	0.319 (0.071-1.027) p=0.082	0.586 (0.164-1.657) p=0.352
	DHEAS	0.414 (0.094-1.296) p=0.172	0.472 (0.133-1.312) p=0.186
	SHBG		
	Progesterone	0.392 (0.089-1.242) p=0.150	0.589 (0.190-1.532) p=0.309
	LH	0.416 (0.095-1.305) p=0.175	0.454 (0.128-1.260) p=0.163
	FSH	0.419 (0.095-1.315) p=0.179	0.465 (0.132-1.292) p=0.177
	ER-α SB	0.398 (0.090-1.249) p=0.154	0.461 (0.130-1.280) p=0.172
	ER-β SB	0.418 (0.095-1.307) p=0.176	0.500 (0.141-1.404) p=0.225
	AR SB	0.274 (0.043-1.004) p=0.091	0.503 (0.142-1.410) p=0.228

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Supplemental Table VI-6: Joint association of high levels of androgens and progesterone (top quantiles) with risk of breast cancer – all cases.

Hormones		Joint association of androgens and progesterone DHEAS Progesterone
		OR* (95% CI)
Not adjusted		1.283 (0.715-1.875) p=0.325
		OR (95% CI)**
Adjusted	Oestradiol	1.278 (0.769-2.101) p=0.337
	Oestrone	1.463 (0.890-2.385) p=0.129
	Androstenedione	0.939 (0.533-1.630) p=0.826
	Testosterone	0.757 (0.421-1.341) p=0.346
	DHEAS	
	SHBG	1.262 (0.760-2.073) p=0.362
	Progesterone	
	LH	1.331 (0.803-2.187) p=0.262
	FSH	1.272 (0.767-2.087) p=0.344
	ER-α SB	1.456 (0.891-2.363) p=0.130
	ER-β SB	1.281 (0.775-2.095) p=0.328
	AR SB	1.381 (0.841-2.248) p=0.197

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Supplemental Table VI-7: Joint association of high levels of androgens and gonadotrophins (top quantiles) with risk of breast cancer – all cases.

Hormones		Joint association of androgens and gonadotrophins		
		Andorestenedione	DHEAS	
		FSH	LH	FSH
		OR* (95% CI)	**OR* (95% CI)**	**OR* (95% CI)**
Not adjusted		2.266 (0.957-5.423) p=0.085	0.991 (0.410-1.832) p=0.982	1.501 (0.501-3.631) p=0.424
		OR (95% CI)**	**OR** (95% CI)**	**OR** (95% CI)**
Adjusted	Oestradiol	2.371 (0.773-4.610) p=0.154	1.012 (0.458-2.122) p=0.976	1.699 (0.629-4.456) p=0.281
	Oestrone	2.748 (0.925-5.578) p=0.072	1.206 (0.553-2.526) p=0.625	1.620 (0.640-3.065) p=0.293
	Androstenedione		0.707 (0.304-1.541) p=0.398	1.306 (0.453-3.570) p=0.606
	Testosterone	1.760 (0.588-3.663) p=0.398	0.676 (0.297-1.457) p=0.330	1.170 (0.403-3.226) p=0.764
	DHEAS	2.341 (0.887-5.100) p=0.087		
	SHBG	2.172 (0.952-5.595) p=0.061	0.888 (0.401-1.859) p=0.759	1.457 (0.511-3.948) p=0.463
	Progesterone	2.368 (0.654-4.113) p=0.273	0.908 (0.398-1.966) p=0.810	1.336 (0.431-3.871) p=0.597
	LH	6.975 (1.141-7.092) **p=0.024**		1.986 (0.675-5.611) p=0.197
	FSH		0.994 (0.450-2.081) p=0.989	
	ER-α SB	2.227 (0.929-5.345) p=0.069	1.000 (0.453-2.089) p=0.999	1.504 (0.490-4.308) p=0.453
	ER-β SB	2.131 (0.892-5.096) p=0.085	1.093 (0.508-2.248) p=0.812	1.342 (0.482-3.509) p=0.554
	AR SB	2.193 (0.915-5.258) p=0.075	0.993 (0.450-2.073) p=0.985	1.149 (0.387-3.110) p=0.790

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Supplemental Table VI-8: Joint association of high levels of androgens and sex hormone-binding globulin (top quantiles) with risk of breast cancer – all cases.

Hormones		Joint association of SHBG and androgens		
		Androstenedione	Testosterone	DHEAS
		SHBG		
		OR* (95% CI)	**OR* (95% CI)**	**OR* (95% CI)**
Not adjusted		0.552 (0.083-1.578) p=0.301	1.085 (0.462-3.475) p=0.875	0.909 (0.278-2.782) p=0.862
		OR (95% CI)**	**OR** (95% CI)**	**OR** (95% CI)**
Adjusted	Oestradiol	0.399 (0.091-1.246) p=0.154	1.237 (0.444-3.228) p=0.669	0.967 (0.296-2.775) p=0.952
	Oestrone	0.448 (0.101-1.419) p=0.121	1.128 (0.382-3.028) p=0.816	0.950 (0.295-2.670) p=0.926
	Androstenedione		0.737 (0.225-2.889) p=0.585	0.772 (0.236-2.199) p=0.642
	Testosterone	0.310 (0.069-0.992) p=0.074		0.656 (0.199-1.891) p=0.454
	DHEAS	0.541 (0.151-1.546) p=0.288	1.064 (0.365-2.824) p=0.906	
	SHBG			
	Progesterone	0.518 (0.144-1.478) p=0.257	1.040 (0.349-2.824) p=0.940	0.832 (0.253-2.399) p=0.744
	LH	0.549 (0.153-1.561) p=0.297	1.104 (0.374-2.963) p=0.848	0.857 (0.266-2.404) p=0.788
	FSH	0.552 (0.154-1.565) p=0.301	1.086 (0.369-2.904) p=0.873	0.907 (0.282-2.540) p=0.859
	ER-α SB	0.557 (0.156-1.583) p=0.309	1.057 (0.358-2.833) p=0.915	0.917 (0.285-2.571) p=0.875
	ER-β SB	0.411 (0.094-1.282) p=0.167	1.091 (0.370-2.924) p=0.866	0.910 (0.283-2.549) p=0.864
	AR SB	0.279 (0.044-1.015) p=0.094	0.901 (0.280-2.521) p=0.849	0.734 (0.201-2.188) p=0.602

*OR values for top-bottom classification were based on controls only and were age adjusted.

**OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Supplemental Table VI-9: Joint association of high levels of progesterone and gonadotrophins (top quantiles) with risk of breast cancer – all cases.

Hormones		Joint association of progesterone and gonadotrophins	
		Progesterone	
		LH	FSH
		OR* (95% CI)	**OR* (95% CI)**
Not adjusted		0.903 (0.0365-1.859) p=0.806	1.133 (0.353-3.277) p=0.826
		OR (95% CI)**	**OR** (95% CI)**
Adjusted	Oestradiol	0.886 (0.374-1.950) p=0.771	1.433 (0.508-3.839) p=0.478
	Oestrone	1.277 (0.565-2.769) p=0.543	1.690 (0.628-4.412) p=0.283
	Androstenedione	0.650 (0.257-1.498) p=0.331	0.677 (0.179-2.117) p=0.524
	Testosterone	0.614 (0.251-1.397) p=0.260	0.888 (0.262-2.698) 0.838
	DHEAS	0.863 (0.347-2.014) p=0.740	1.116 (0.331-3.372) p=0.849
	SHBG	0.816 (0.344-1.803) p=0.627	1.059 (0.318-3.155) p=0.920
	Progesterone		
	LH		1.323 (0.393-4.004) p=0.629
	FSH	0.908 (0.383-1.998) p=0.816	
	ER-α SB	1.145 (0.514-2.431) p=0.730	1.480 (0.483-4.232) p=0.470
	ER-β SB	0.903 (0.381-1.989) p=0.807	1.134 (0.342-3.362) p=0.825
	AR SB	1.128 (0.506-2.391) p=0.759	1.251 (0.416-3.357) p=0.673

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Supplemental Table VI-10: Joint association of high levels of gonadotrophins (top quantiles) with risk of breast cancer – all cases.

Hormones		Joint association of gonadotrophins
		LH
		FSH
		OR* (95% CI)
Not adjusted		0.851 (0.476-1.513) p=0.589
		OR (95% CI)**
Adjusted	Oestradiol	0.941 (0.516-1.664) p=0.838
	Oestrone	0.918 (0.504-1.620) p=0.773
	Androstenedione	0.787 (0.415-1.433) p=0.447
	Testosterone	0.879 (0.474-1.576) p=0.673
	DHEAS	0.852 (0.464-1.510) p=0.593
	SHBG	0.900 (0.487-1.608) p=0.728
	Progesterone	0.837 (0.448-1.504) p=0.562
	LH	
	FSH	
	ER-α SB	0.843 (0.452-1.514) p=0.578
	ER-β SB	0.798 (0.429-1.428) p=0.460
	AR SB	0.804 (0.432-1.440) p=0.476

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Supplemental Table VI-11: Joint association of high levels of sex hormone-binding globulin and progesterone (top quantiles) with risk of breast cancer – all cases.

Hormones		Joint association of SHBG and progesterone SHBG Progesterone
		OR* (95% CI)
Not adjusted		0.555 (0.142-1.412) p=0.306
		OR (95% CI)**
Adjusted	Oestradiol	0.492 (0.137-1.396) p=0.218
	Oestrone	0.605 (0.195-1.576) p=0.335
	Androstenedione	0.463 (0.128-1.342) p=0.188
	Testosterone	0.416 (0.114-1.215) p=0.136
	DHEAS	0.531 (0.146-1.542) p=0.280
	SHBG	
	Progesterone	
	LH	0.519 (0.145-1.479) p=0.256
	FSH	0.546 (0.153-1.553) p=0.293
	ER-α SB	0.635 (0.204-1.671) p=0.387
	ER-β SB	0.515 (0.145-1.444) p=0.245
	AR SB	0.481 (0.136-1.334) p=0.196

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Supplemental Table VI-12: Joint association of high levels of sex hormone-binding globulin and gonadotrophins (top quantiles) with risk of breast cancer – all cases.

Hormones		Joint association of SHBG and gonadotrophins	
		SHBG	
		LH	FSH
		OR* (95% CI)	**OR* (95% CI)**
Not adjusted		0.229 (0.052-0.892) p=0.051	0.580 (0.278-1.785) p=0.250
		OR (95% CI)**	**OR** (95% CI)**
Adjusted	Oestradiol	0.224 (0.035-0.796) **p=0.047**	0.716 (0.275-1.668) p=0.460
	Oestrone	0.235 (0.037-0.835) p=0.055	0.601 (0.215-1.453) p=0.287
	Androstenedione	0.120 (0.007-0.597) **p=0.041**	0.596 (0.213-1.438) p=0.278
	Testosterone	0.246 (0.038-0.894) p=0.066	0.535 (0.174-1.370) p=0.225
	DHEAS	0.230 (0.036-0.817) p=0.051	0.581 (0.209-1.396) p=0.254
	SHBG		
	Progesterone	0.229 (0.036-0.815) p=0.051	0.457 (0.150-1.148) p=0.123
	LH		0.664 (0.236-1.625) p=0.397
	FSH	0.217 (0.043-0.790) **p=0.046**	
	ER-α SB	0.243 (0.038-0.871) p=0.062	0.493 (0.162-1.243) p=0.165
	ER-β SB	0.229 (0.036-0.814) p=0.051	0.479 (0.157-1.210) p=0.149
	AR SB	0.230 (0.036-0.815) p=0.051	0.487 (0.160-1.225) p=0.157

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Supplemental Table VI-13: Joint association of high levels of oestrogens and androgens (top quantiles) with risk of breast cancer – cases that gave a sample less than 2 years before diagnosis.

Hormones		Joint association of oestrogens and androgens			
		Oestradiol			Oestrone
		Androstenedione	Testosterone	DHEAS	DHEAS
		OR* (95% CI)	**OR* (95% CI)**	**OR* (95% CI)**	**OR* (95% CI)**
Not adjusted		1.701 (0.818-4.014) p=0.179	1.594 (0.872-3.527) p=0.199	0.870 (0.303-2.266) p=0.786	1.578 (0.696-4.398) p=0.330
		OR (95% CI)**	**OR** (95% CI)**	**OR** (95% CI)**	**OR** (95% CI)**
Adjusted	Oestradiol				1.324 (0.463-3.322) p=0.570
	Oestrone	1.618 (0.710-3.446) p=0.227	1.514 (0.715-3.032) p=0.256	0.834 (0.270-2.137) p=0.726	
	Androstenedione		1.287 (0.581-2.706) p=0.517	0.731 (0.231-1.933) p=0.556	1.078 (0.367-2.792) p=0.882
	Testosterone	1.086 (0.448-2.483) p=0.850		0.518 (0.162-1.385) p=0.220	0.903 (0.305-2.353) p=0.843
	DHEAS	1.680 (0.730-3.626) p=0.200	1.582 (0.736-3.229) p=0.220		
	SHBG	1.662 (0.705-3.689) p=0.224	1.503 (0.700-3.062) p=0.275	0.740 (0.233-1.977) p=0.575	1.437 (0.515-3.660) p=0.463
	Progesterone	1.669 (0.731-3.657) p=0.201	1.592 (0.751-3.196) p=0.204	0.875 (0.284-2.236) p=0.795	1.395 (0.519-3.390) p=0.480
	LH	1.671 (0.737-3.544) p=0.195	1.579 (0.749-3.146) p=0.208	0.859 (0.280-2.183) p=0.768	1.574 (0.588-3.803) p=0.334
	FSH	1.709 (0.753-3.631) p=0.177	1.639 (0.774-3.284) p=0.177	0.894 (0.290-2.283) p=0.828	1.621 (0.605-3.929) p=0.304
	ER-α SB	1.592 (0.676-3.464) p=0.259	1.449 (0.671-2.934) p=0.32	0.868 (0.283-2.205) p=0.875	1.579 (0.591-3.812) p=0.330
	ER-β SB	1.528 (0.650-3.315) p=0.302	1.446 (0.668-2.934) p=0.324	0.795 (0.284-2.231) p=0.784	1.854 (0.728-4.362) p=0.171
	AR SB	1.531 (0.652-3.314) p=0.299	1.452 (0.672-2.940) p=0.317	0.871 (0.672-2.211) p=0.788	1.581 (0.590-3.824) p=0.330

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Supplemental Table VI-14: Joint association of high levels of oestrogens and progesterone (top quantiles) with risk of breast cancer – cases that gave a sample less than 2 years before diagnosis.

Hormones		Joint association of oestrogens and progesterone	
		Progesterone	
		Oestradiol	Oestrone
		OR* (95% CI)	**OR* (95% CI)**
Not adjusted		1.181 (0.475-2.813) p=0.714	2.015 (0.897-5.412) p=0.123
		OR (95% CI)**	**OR** (95% CI)**
Adjusted	Oestradiol		2.161 (0.838-5.212) p=0.094
	Oestrone	1.156 (0.440-2.702) p=0.751	
	Androstenedione	0.939 (0.344-2.306) p=0.896	1.576 (0.591-3.921) p=0.341
	Testosterone	0.802 (0.294-1.963) p=0.645	1.465 (0.554-3.605) p=0.419
	DHEAS	1.144 (0.425-2.765) p=0.775	1.996 (0.739-5.079) p=0.155
	SHBG		
	Progesterone	1.334 (0.506-3.153) p=0.531	1.948 (0.756-4.678) p=0.146
	LH	1.158 (0.443-2.691) p=0.747	1.979 (0.773-4.713) p=0.134
	FSH	1.217 (0.464-2.848) p=0.667	2.090 (0.814-5.000) p=0.107
	ER-α SB	1.001 (0.358-2.419) p=0.999	1.752 (0.650-4.296) p=0.237
	ER-β SB	1.010 (0.361-2.442) p=0.983	1.857 (0.685-4.591) p=0.195
	AR SB	0.998 (0.357-2.412) p=0.999	1.744 (0.646-4.278) p=0.242

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Supplemental Table VI-15: Joint association of high levels of oestrogens and gonadotrophins (top quantiles) with risk of breast cancer – cases that gave a sample less than 2 years before diagnosis.

Hormones		Joint association of oestrogens and gonadotrophins			
		Oestradiol		Oestrone	
		LH	FSH	LH	FSH
		OR* (95% CI)	**OR* (95% CI)**	**OR* (95% CI)**	**OR* (95% CI)**
Not adjusted		1.047 (0.258-3.504) p=0.945	1.422 (0.365-5.493) p=0.609	1.488 (0.541-4.603) p=0.461	2.273 (0.980-6.940) p=0.095
		OR (95% CI)**	**OR** (95% CI)**	**OR** (95% CI)**	**OR** (95% CI)**
Adjusted	Oestradiol			1.465 (0.060-3.998) p=0.480	1.898 (0.645-5.035) p=0.213
	Oestrone	1.134 (0.249-3.386) p=0.851	1.556 (0.328-5.757) p=0.530		
	Androstenedione	0.997 (0.220-3.315) p=0.996	1.396 (0.298-4.995) p=0.630	1.260 (0.470-3.687) p=0.694	1.837 (0.619-4.920) p=0.242
	Testosterone	0.837 (0.183-2.817) p=0.792	1.352 (0.287-4.884) p=0.665	1.315 (0.347-3.660) p=0.619	1.858 (0.624-4.990) p=0.234
	DHEAS	1.031 (0.229-3.397) p=0.963	1.434 (0.309-5.081) p=0.601	1.464 (0.421-4.002) p=0.481	2.286 (0.829-5.854) p=0.092
	SHBG	1.003 (0.223-3.307) p=0.996	1.436 (0.309-5.097) p=0.599	1.467 (0.041-4.008) p=0.477	2.295 (0.832-5.890) p=0.091
	Progesterone	1.150 (0.253-3.876) p=0.835	1.327 (0.284-4.750) p=0.683	1.412 (0.441-3.895) p=0.526	2.590 (0.920-6.866) p=0.059
	LH		1.635 (0.344-6.036) p=0.486		2.427 (0.874-6.292) p=0.074
	FSH	1.002 (0.221-3.333) p=0.998		1.444 (0.451-3.960) p=0.498	
	ER-α SB	1.049 (0.234-3.446) p=0.942	1.425 (0.307-5.045) p=0.607	1.504 (0.471-4.117) p=0.451	2.281 (0.828-5.841) p=0.093
	ER-β SB	1.062 (0.236-3.501) p=0.928	1.427 (0.307-5.070) p=0.606	1.522 (0.469-4.168) p=0.437	2.265 (0.820-5.814) p=0.096
	AR SB	1.050 (0.234-3.447) p=0.942	1.434 (0.309-5.076) p=0.601	1.495 ().469-4.085) p=0.456	2.289 (0.831-5.865) p=0.091

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Supplemental Table VI-16: Joint association of high levels of oestrogens and sex hormone-binding globulin (top quantiles) with risk of breast cancer – cases that gave a sample less than 2 years before diagnosis.

Hormones		Joint association of SHBG and oestrogens	
		Oestradiol	Oestrone
		SHBG	
		OR* (95% CI)	**OR* (95% CI)**
Not adjusted		0.267 (0.033-1.457) p=0.206	0.456 (0.039-1.978) p=0.301
		OR (95% CI)**	**OR** (95% CI)**
Adjusted	Oestradiol		0.477 (0.074-1.732) p=0.331
	Oestrone	0.275 (0.014-1.603) p=0.217	
	Androstenedione	0.257 (0.015-1.569) p=0.196	0.526 (0.081-1.929) p=0.402
	Testosterone	0.213 (0.012-1.485) p=0.142	0.565 (0.088-2.069) p=0.455
	DHEAS	0.265 (0.014-1.547) p=0.204	0.465 (0.073-1.677) p=0.314
	SHBG		
	Progesterone	0.237 (0.013-1.523) p=0.171	0.690 (0.158-2.126) p=0.562
	LH	0.268 (0.015-1.492) p=0.207	0.448 (0.070-1.613) p=0.290
	FSH	0.272 (0.015-1.504) p=0.213	0.448 (0.070-1.615) p=0.290
	ER-α SB	0.268 (0.015-1.587) p=0.207	0.456 (0.071-1.642) p=0.301
	ER-β SB	0.261 (0.014-1.511) p=0.199	0.452 (0.070-1.648) p=0.298
	AR SB	0.267 (0.015-1.563) p=0.206	0.490 (0.076-1.779) p=0.349

*OR values for top-bottom classification were based on controls only and were age adjusted.

**OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Supplemental Table VI-17: Joint association of high levels of androgens (top quantiles) with risk of breast cancer – cases that gave a sample less than 2 years before diagnosis.

Hormones		Joint association of androgens
		Testosterone
		DHEAS
		OR* (95% CI)
Not adjusted		1.617 (0.921-3.210) p=0.131
		OR (95% CI)**
Adjusted	Oestradiol	1.679 (0.873-3.125) p=0.109
	Oestrone	1.588 (0.826-2.948) p=0.152
	Androstenedione	
	Testosterone	1.215 (0.577-2.472) p=0.599
	DHEAS	
	SHBG	1.550 (0.809-2.872) p=0.172
	Progesterone	1.592 (0.724-3.473) p=0.242
	LH	1.657 (0.867-3.061) p=0.114
	FSH	1.628 (0.852-3.008) p=0.128
	ER-α SB	1.659 (0.870-3.056) p=0.112
	ER-β SB	1.624 (0.850-3.000) p=0.130
	AR SB	1.581 (0.830-2.907) p=0.150

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Supplemental Table VI-18: Joint association of high levels of androgens and progesterone (top quantiles) with risk of breast cancer – cases that gave a sample less than 2 years before diagnosis.

Hormones		Joint association of androgens and progesterone DHEAS Progesterone
		OR* (95% CI)
Not adjusted		1.532 (0.753-2.414) p=0.167
		OR (95% CI)**
Adjusted	Oestradiol	1.580 (0.844-2.878) p=0.142
	Oestrone	1.537 (0.818-2.807) p=0.170
	Androstenedione	1.188 (0.591-2.315) p=0.620
	Testosterone	0.974 (0.479-1.929) p=0.942
	DHEAS SHBG Progesterone	1.484 (0.792-2.705) p=0.205
	LH	1.577 (0.844-2.867) p=0.142
	FSH	1.528 (0.817-2.777) p=0.172
	ER-α SB	1.532 (0.822-2.773) p=0.167
	ER-β SB	1.543 (0.828-2.794) p=0.160
	AR SB	1.533 (0.823-2.776) p=0.166

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Supplemental Table VI-19: Joint association of high levels of androgens and gonadotrophins (top quantiles) with risk of breast cancer – cases that gave a sample less than 2 years before diagnosis.

Hormones		Joint association of androgens and gonadotrophins			
		Androstenedione		DHEAS	
		LH	FSH	LH	FSH
		OR* (95% CI)	**OR* (95% CI)**	**OR* (95% CI)**	**OR* (95% CI)**
	Not adjusted	2.429 (0.854-5.197) p=0.057	1.890 (0.703-6.343) p=0.250	1.234 (0.455-2.679) p=0.646	1.634 (0.472-5.124) p=0.421
		OR (95% CI)**	**OR** (95% CI)**	**OR** (95% CI)**	**OR** (95% CI)**
Adjusted	Oestradiol	2.583 (0.981-6.410) **p=0.045**	1.473 (0.400-4.442) p=0.517	1.269 (0.481-2.993) p=0.604	2.013 (0.606-5.926) p=0.219
	Oestrone	3.059 (1.139-7.816) **p=0.021**	1.336 (0.560-5.529) p=0.273	1.604 (0.632-3.746) p=0.292	2.158 (0.720-5.862) p=0.143
	Androstenedione			0.859 (0.296-2.164) p=0.760	1.470 (0.385-4.679) p=0.535
	Testosterone	1.917 (0.702-4.940) p=0.186	1.146 (0.304-3.548) p=0.824	0.879 (0.324-2.138) p=0.787	1.349 (0.352-4.323) p=0.631
	DHEAS	2.492 (0.916-6.428) p=0.063	1.864 (0.570-5.327) p=0.264		
	Progesterone	2.242 (0.854-5.536) p=0.086	2.079 (0.628-6.081) p=0.197	1.083 (0.410-2.553) p=0.863	1.594 (0.421-5.038) p=0.450
	SHBG	2.441 (0.907-6.214) p=0.066	1.036 (0.229-3.432) p=0.958	1.078 (0.391-3.684) p=0.877	1.203 (0.257-4.272) p=0.789
	LH		2.291 (0.687-6.791) p=0.147		1.985 (0.507-6.607) p=0.283
	FSH	2.429 (0.926-5.979) p=0.059		1.224 (0.464-2.884) p=0.661	
	ER-α SB	2.617 (0.996-6.487) **p=0.041**	1.881 (0.575-5.382) p=0.257	1.235 (0.470-2.898) p=0.645	1.344 (0.291-5.038) p=0.665
	ER-β SB	2.390 (0.914-5.864) p=0.063	1.855 (0.569-5.284) p=0.266	1.408 (0.563-3.216) p=0.436	1.441 (0.388-4.384) p=0.544
	AR SB	2.449 (0.937-6.006) p=0.056	1.928 (0.590-5.512) p=0.239	1.234 (0.469-2.898) p=0.647	1.080 (0.239-3.582) p=0.908

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Supplemental Table VI-20: Joint association of high levels of androgens and sex hormone-binding globulin (top quantiles) with risk of breast cancer – cases that gave a sample less than 2 years before diagnosis.

Hormones		Joint association of SHBG and androgens		
		Androstenedione	Testosterone	DHEAS
		SHBG		
		OR* (95% CI)	**OR* (95% CI)**	**OR* (95% CI)**
Not adjusted		0.276 (0.028-) p=0.217	0.703 (0.054-2.873) p=0.650	0.709 (0.277-3.091) p=0.659
		OR (95% CI)**	**OR** (95% CI)**	**OR** (95% CI)**
Adjusted	Oestradiol	not enough points	0.716 (0.108-2.781) p=0.670	0.787 (0.166-2.961) p=0.761
	Oestrone	not enough points	0.755 (0.115-2.930) p=0.719	0.766 (0.096-2.481) p=0.733
	Androstenedione		0.300 (0.016-1.609) p=0.256	0.635 (0.077-2.032) p=0.564
	Testosterone	not enough points		0.513 (0.093-2.632) p=0.400
	DHEAS	0.269 (0.008-1.373) p=0.209	0.666 (0.101-2.594) p=0.606	
	SHBG			
	Progesterone	0.251 (0.014-1.289) p=0.187	0.625 (0.094-2.439) p=0.551	0.604 (0.090-2.406) p=0.527
	LH	0.267 (0.015-1.357) p=0.205	0.703 (0.107-2.687) p=0.651	0.677 (0.103-2.609) p=0.618
	FSH	0.275 (0.015-1.396) p=0.215	0.703 (0.107-2.686) p=0.651	0.712 (0.108-2.734) p=0.663
	ER-α SB	0.275 (0.015-1.398) p=0.216	0.705 (0.108-2.698) p=0.654	0.710 (0.108-2.728) p=0.661
	ER-β SB	not enough points	0.654 (0.100-2.508) p=0.587	0.679 (0.103-2.613) p=0.621
	AR SB	not enough points	0.703 (0.107-2.686) p=0.651	0.708 (0.108-2.723) p=0.659

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Supplemental Table VI-21: Joint association of high levels of gonadotrophins and progesterone (top quantiles) with risk of breast cancer – cases that gave a sample less than 2 years before diagnosis.

Hormones		Joint association of progesterone and gonadotrophins	
		Progesterone	
		LH	FSH
		OR* (95% CI)	**OR* (95% CI)**
Not adjusted		1.396 p=0.469	1.304 p=0.696
		OR* (95% CI)	**OR* (95% CI)**
Adjusted	Oestradiol	1.380 (0.522-3.274) p=0.486	1.964 (0.593-5.751) p=0.234
	Oestrone	1.628 (0.606-3.967) p=0.302	2.120 (0.639-6.229) p=0.186
	Androstenedione	1.012 (0.350-2.557) p=0.980	0.678 (0.101-2.716) p=0.626
	Testosterone	0.980 (0.358-2.414) p=0.967	1.064 (0.227-3.784) p=0.920
	SHBG	1.236 (0.466-2.943) p=0.648	1.206 (0.261-4.206) p=0.784
	DHEAS	1.382 (0.486-3.583) p=0.518	1.264 (0.271-4.473) p=0.734
	Progesterone		
	LH		1.444 (0.310-5.115) p=0.595
	FSH	1.402 (0.530-3.324) p=0.464	
	ER-α	1.397 (0.528-3.311) p=0.469	1.295 (0.281-4.505) p=0.705
	ER-β	1.381 (0.521-3.279) p=0.486	1.255 (0.2.72-4.348) p=0.739
	AR	1.400 (0.529-3.322) p=0.467	1.167 (0.256-3.941) p=0.818

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydro- epiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Supplemental Table VI-22: Joint association of high levels of gonadotrophins (top quantiles) with risk of breast cancer – cases that gave a sample less than 2 years before diagnosis.

Hormones		Joint association of gonadotrophins
		LH
		FSH
		OR* (95% CI)
Not adjusted		0.920 p=0.823
		OR* (95% CI)
Adjusted	Oestradiol	1.059 (0.496-2.098) p=0.876
	Oestrone	1.077 (0.504-2.141) p=0.840
	Androstenedione	0.793 (0.332-1.689) p=0.571
	Testosterone	0.975 (0.441-1.981) p=0.976
	SHBG	0.986 (0.0447-2.002) p=0.969
	DHEAS	0.923 (0.421-1.857) p=0.832
	Progesterone	0.869 (0.381-1.795) p=0.719
	LH	
	FSH	
	ER-α	0.864 (0.380-1.780) p=0.707
	ER-β	0.795 (0.349-1.638) p=0.556
	AR	0.818 (0.360-1.684) p=0.607

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydro- epiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Supplemental Table VI-23: Joint association of high levels of progesterone and sex hormone-binding globulin (top quantiles) with risk of breast cancer – cases that gave a sample less than 2 years before diagnosis.

Hormones		Joint association of SHBG and progesterone
		SHBG Progesterone
		OR* (95% CI)
Not adjusted		0.539 (0.082-1.916) p=0.420
		OR (95% CI)**
Adjusted	Oestradiol	0.511 (0.079-1.895) p=0.384
	Oestrone	0.474 (0.074-1.716) p=0.327
	Androstenedione	0.480 (0.074-1.797) p=0.343
	Testosterone	0.421 (0.064-1.592) p=0.266
	DHEAS	0.511 (0.078-1.920) p=0.386
	SHBG	
	Progesterone	
	LH	0.514 (0.079-1.897) p=0.386
	FSH	0.533 (0.082-1.966) p=0.412
	ER-α SB	0.502 (0.078-1.827) p=0.366
	ER-β SB	0.487 (0.076-1.776) p=0.346
	AR SB	0.468 (0.073-1.687) p=0.318

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydro- epiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Supplemental Table VI-24: Joint association of high levels of sex hormone-binding globulin and gonadotrophins (top quantiles) with risk of breast cancer – cases that gave a sample less than 2 years before diagnosis.

Hormones		Joint association of SHBG and gonadotrophins	
		SHBG	
		LH	FSH
		OR* (95% CI)	**OR* (95% CI)**
Not adjusted		0.221 (0.013-1.121) p=0.145	0.565 (0.142-1.590) p=0.364
		OR (95% CI)**	**OR** (95% CI)**
Adjusted	Oestradiol	0.218 (0.012-1.089) p=0.142	0.777 (0.221-2.135) p=0.655
	Oestrone	0.232 (0.013-1.163) p=0.159	0.577 (0.133-1.749) p=0.386
	Androstenedione	not enough points	0.597 (0.138-1.801) p=0.415
	Testosterone	0.250 (0.014-1.273) p=0.184	0.440 (0.069-1.570) p=0.278
	DHEAS	0.223 (0.012-1.122) p=0.149	0.570 (0.057-1.713) p=0.374
	SHBG		
	Progesterone	0.226 (0.027-1.136) p=0.153	0.362 (0.140-1.268) p=0.175
	LH		0.611 (0.103-1.881) p=0.442
	FSH	0.200 (0.012-1.032) p=0.125	
	ER-α SB	0.236 (0.013-1.182) p=0.164	0.373 (0.059-1.309) p=0.189
	ER-β SB	0.209 (0.011-1.047) p=0.132	0.344 (0.054-1.211) p=0.156
	AR SB	0.221 (0.012-1.103) p=0.146	0.373 (0.059-1.310) p=0.189

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydro- epiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Supplemental Table VI-25: Joint association of high levels of oestrogens (top quantiles) with risk of breast cancer – cases that gave a sample more than 2 years before diagnosis.

Hormones		Joint association of oestrogens Oestradiol Oestrone
		OR* (95% CI)
Not adjusted		1.437 (0.665-4.908) p=0.497
		OR (95% CI)**
Adjusted	Oestradiol	
	Oestrone	
	Androstenedione	1.340 (0.420-3.643) p=0.587
	Testosterone	1.232 (0.383-3.381) p=0.701
	DHEAS	1.441 (0.455-3.891) p=0.495
	SHBG	1.347 (0.425-3.644) p=0.579
	Progesterone	1.422 (0.499-3.831) p=0.511
	LH	1.369 (0.432-3.703) p=0.558
	FSH	1.434 (0.453-3.870) p=0.501
	ER-α	1.397 (0.439-3.792) p=0.535
	ER-β	1.404 (0.443-3.796) p=0.527
	AR	1.441 (0.455-3.884) p=0.495

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydro- epiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Supplemental Table VI-26: Joint association of high levels of oestrogens and androgens (top quantiles) with risk of breast cancer – cases that gave a sample more than 2 years before diagnosis.

Hormones		Joint association of oestrogens and androgens			
		Oestradiol	Oestrone		
		DHEAS	Androstenedione	Testosterone	DHEAS
		OR* (95% CI)	**OR* (95% CI)**	**OR* (95% CI)**	**OR* (95% CI)**
Not adjusted		1.331 (0.515-3.069) p=0.528	1.202 (0.502-4.037) p=0.725	1.669 (0.773-4.910) p=0.272	0.942 (0.294-2.757) p=0.917
		OR (95% CI)**	**OR** (95% CI)**	**OR** (95% CI)**	**OR** (95% CI)**
Adjusted	Oestradiol		1.221 (0.389-3.232) p=0.706	1.680 (0.626-4.085) p=0.271	0.890 (0.249-2.503) p=0.839
	Oestrone	1.318 (0.504-3.075) p=0.543			
	Androstenedione	1.079 (0.400-2.610) p=0.873		1.471 (0.602-3.637) p=0.422	0.750 (0.214-2.175) p=0.622
	Testosterone	0.810 (0.295-2.005) p=0.663	1.002 (0.310-2.757) p=0.997		0.602 (0.164-1.762) p=0.391
	DHEAS		1.208 (0.382-3.236) p=0.723	1.740 (0.634-4.352) p=0.253	
	Progesterone	1.351 (0.516-3.153) p=0.508	1.149 (0.368-3.021) p=0.791	1.543 (0.578-3.726) p=0.354	0.873 (0.244-2.461) p=0.813
	SHBG	1.425 (0.522-3.525) p=0.461	1.254 (0.388-3.465) p=0.679	1.681 (0.622-4.118) p=0.275	0.829 (0.225-2.455) p=0.753
	LH	1.300 (0.496-3.037) p=0.564	1.124 (0.359-2.957) p=0.824	1.609 (0.601-3.885) p=0.311	0.929 (0.260-2.623) p=0.898
	FSH	1.328 (0.506-3.111) p=0.535	1.203 (0.386-3.152) p=0.725	1.667 (0.623-4.027) p=0.276	0.937 (0.263-2.636) p=0.910
	ER-α SB	1.385 (0.528-3.243) p=0.475	1.201 (0.384-3.165) p=0.728	1.710 (0.638-4.138) p=0.254	0.944 (0.264-2.666) p=0.920
	ER-β SB	1.347 (0.513-3.54) p=0.514	1.172 (0.375-3.078) p=0.763	1.648 (0.617-3.971) p=0.286	0.932 (0.261-2.626) p=0.903
	AR SB	1.122 (0.400-2.717) p=0.811	1.222 (0.391-3.207) p=0.703	1.676 (0.629-4.034) p=0.269	0.910 (0.255-2.559) p=0.869

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydro- epiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Supplemental Table VI-27: Joint association of high levels of oestrogens and progesterone (top quantiles) with risk of breast cancer – cases that gave a sample more than 2 years before diagnosis.

Hormones		Joint association of oestrogens and progesterone	
		Oestradiol	Oestrone
		Progesterone	
		OR* (95% CI)	**OR* (95% CI)**
Not adjusted		1.067 (0.418-2.753) p=0.893	1.297 (0.451-3.644) p=0.624
		OR (95% CI)**	**OR** (95% CI)**
Adjusted	Oestradiol		1.073 (0.298-3.074) p=0.903
	Oestrone	1.254 (0.480-2.921) p=0.618	
	Androstenedione	0.787 (0.271-1.999) p=0.635	0.986 (0.355-2.697) p=0.979
	Testosterone	0.718 (0.245-1.834) p=0.512	0.506 (0.113-1.635) p=0.303
	DHEAS	1.090 (0.376-2.767) p=0.863	1.353 (0.409-3.876) p=0.591
	SHBG	1.147 (0.408-2.790) p=0.776	1.240 (0.394-3.302) 0.686
	Progesterone		
	LH	1.009 (0.360-2.445) p=0.985	1.274 (0.404-3.397) p=0.649
	FSH	1.057 (0.376-2.116) p=0.908	1.292 (0.411-4.33) p=0.629
	ER-α	1.290 (0.492-3.020) p=0.576	1.290 (0.409-3.444) p=0.632
	ER-β	1.069 (0.381-2.584) p=0.890	1.233 (0.391-3.288) p=0.694
	AR	1.053 (0.375-2.551) p=0.914	1.274 (0.405-3.386) p=0.647

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydro- epiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Supplemental Table VI-28: Joint association of high levels of oestrogens and gonadotrophins (top quantiles) with risk of breast cancer – cases that gave a sample more than 2 years before diagnosis.

Hormones		Joint association of oestrogens and gonadotrophins			
		Oestradiol		Oestrone	
		LH	FSH	LH	FSH
		OR* (95% CI)	**OR* (95% CI)**	**OR* (95% CI)**	**OR* (95% CI)**
Not adjusted		0.363 (0.044-2.788) p=0.335	Not enough points	0.617 (0.124-2.560) p=0.530	0.650 (0.148-3.127) p=0.577
		OR (95% CI)**	**OR** (95% CI)**	**OR** (95% CI)**	**OR** (95% CI)**
Adjusted	Oestradiol			0.618 (0.095-2.296) p=0.531	0.657 (0.101-2.470) p=0.587
	Oestrone	0.396 (0.021-2.118) p=0.381			
	Androstenedione	0.329 (0.018-1.750) p=0.293		0.648 (0.078-2.422) p=0.574	0.579 (0.088-2.205) p=0.484
	Testosterone	0.284 (0.015-1.529) p=0.235		0.593 (0.091-2.239) p=0.501	0.706 (0.108-2.674) p=0.654
	DHEAS	0.362 (0.020-1.909) p=0.334		0.617 (0.095-2.296) p=0.531	0.650 (0.100-2.441) p=0.577
	SHBG	0.363 (0.021-2.043) p=0.366	Not enough points	0.617 (0.092-2.236) p=0.507	0.650 (0.111-2.764) p=0.681
	Progesterone	0.361 (0.020-1.900) p=0.333		0.603 (0.093-2.241) p=0.511	0.642 (0.099-2.409) p=0.566
	LH				0.747 (0.114-2.865) p=0.708
	FSH	0.363 (0.020-2.568) p=0.336		0.620 (0.096-2.316) p=0.536	
	ER-α	0.354 (0.019-1.872) p=0.324		0.548 (0.084-2.059) p=0.438	0.615 (0.094-2.331) p=0.532
	ER-β	0.340 (0.018-1.798) p=0.306		0.616 (0.095-2.296) p=0.529	0.667 (0.102-2.515) p=0.602
	AR	0.350 (0.019-1.842) p=0.318		0.639 (0.099-2.380) p=0.561	0.678 (1.04-2.555) p=0.616

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Supplemental Table VI-29: Joint association of high levels of oestrogens and sex hormone-binding globulin (top quantiles) with risk of breast cancer – cases that gave a sample more than 2 years before diagnosis.

Hormones		Joint association of oestrogens and SHBG	
		Oestradiol	Oestrone
		SHBG	
		OR* (95% CI)	**OR* (95% CI)**
Not adjusted		0.267 (0.033-1.457) p=0.206	0.456 (0.039-1.978) p=0.301
		OR (95% CI)**	**OR** (95% CI)**
Adjusted	Oestradiol		0.477 (0.074-1.732) p=0.331
	Oestrone	0.275 (0.014-1.498) p=0.217	
	Androstenedione	0.257 (0.015-1.465) p=0.196	0.526 (0.081-1.929) p=0.402
	Testosterone	0.213 (0.012-1.758) p=0.142	0.565 (0.088-2.069) p=0.455
	DHEAS	0.265 (0.014-1.498) p=0.204	0.465 (0.073-1.677) p=0.314
	Progesterone		
	SHBG	0.237 (0.013-1.432) p=0.171	0.690 (0.158-2.126) p=0.562
	LH	0.268 (0.015-1.785) p=0.207	0.448 (0.070-1.613) p=0.290
	FSH	0.272 (0.015-1.432) p=0.213	0.448 (0.070-1.615) p=0.290
	ER-α SB	0.268 (0.015-1.527) p=0.207	0.456 (0.071-1.642) p=0.301
	ER-β SB	0.261 (0.014-1.791) p=0.199	0.452 (0.070-1.648) p=0.298
	AR SB	0.267 (0.015-1.451) p=0.206	0.490 (0.076-1.779) p=0.349

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Supplemental Table VI-30: Joint association of high levels of androgens (top quantiles) with risk of breast cancer – cases that gave a sample more than 2 years before diagnosis.

Hormones		Joint association of androgens	
		Androstenedione DHEAS	Testosterone DHEAS
		OR* (95% CI)	**OR* (95% CI)**
Not adjusted		1.897 (0.884-3.458) p=0.067	1.611 (0.870-3.108) p=0.141
		OR (95% CI)**	**OR** (95% CI)**
Adjusted	Oestradiol	1.888 (0.917-3.735) p=0.074	1.552 (0.794-2.923) p=0.184
	Oestrone	2.000 (0.976-3.942) **p=0.050**	1.630 (0.840-3.048) p=0.135
	Androstenedione		1.207 (0.575-2.447) p=0.609
	Testosterone	1.103 (0.484-2.455) p=0.813	
	DHEAS		
	Progesterone	1.897 (0.929-3.720) p=0.068	1.587 (0.819-2.960) p=0.156
	SHBG	2.801 (1.141-7.047) **p=0.025**	1.920 (0.847-4.315) p=0.114
	LH	1.993 (0.974-3.924) p=0.051	1.662 (0.857-3.106) p=0.120
	FSH	1.895 (0.929-3.707) p=0.068	1.598 (0.824-2.983) p=0.151
	ER-α SB	1.967 (0.960-3.868) p=0.055	1.647 (0.848-3.079) p=0.127
	ER-β SB	1.925 (0.939-3.792) p=0.064	1.611 (0.830-3.010) p=0.145
	AR SB	1.718 (0.826-3.410) p=0.132	1.437 (0.731-2.704) p=0.274

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Supplemental Table VI-31: Joint association of high levels of androgens and progesterone (top quantiles) with risk of breast cancer – cases that gave a sample more than 2 years before diagnosis.

Hormones		Joint association of androgens and progesterone		
		Androstenedione	Testosterone	DHEAS
		Progesterone		
		OR* (95% CI)	**OR* (95% CI)**	**OR* (95% CI)**
Not adjusted		1.535 (0.751-2.988) p=0.224	1.618 (0.827-3.052) p=0.148	1.038 (0.497-1.857) p=0.915
		OR (95% CI)**	**OR** (95% CI)**	**OR** (95% CI)**
Adjusted	Oestradiol	1.470 (0.706-2.912) p=0.283	1.569 (0.788-2.996) p=0.183	0.994 (0.486-1.920) p=0.987
	Oestrone	1.646 (0.814-3.186) p=0.149	1.884 (0.965-3.555) p=0.055	1.384 (0.714-2.580) p=0.319
	Androstenedione		1.148 (0.531-2.394) p=0.719	0.749 (0.346-1.531) p=0.443
	Testosterone	0.893 (0.394-1.943) p=0.780		0.597 (0.270-1.254) p=0.186
	DHEAS	1.869 (0.795-4.337) p=0.145	2.191 (0.932-5.198) p=0.071	
	SHBG	1.554 (0.752-3.055) p=0.215	1.613 (0.817-3.054) p=0.153	1.033 (0.507-2.560) p=0.926
	Progesterone			
	LH	1.515 (0.733-2.973) p=0.242	1.638 (0.829-3.104) p=0.140	1.065 (0.522-2.052) p=0.856
	FSH	1.522 (0.738-2.981) p=0.235	1.605 (0.813-3.040) p=0.157	1.023 (0.502-1.969) p=0.948
	ER-α	1.624 (0.803-3.142) p=0.161	1.640 (0.828-3.115) p=0.141	1.375 (0.710-2.560) p=0.328
	ER-β	1.515 (0.735-2.969) p=0.240	1.634 (0.827-3.097) p=0.143	1.021 (0.501-1.962) p=0.952
	AR	1.513 (0.733-2.969) p=0.242	1.695 (0.855-3.226) p=0.117	1.223 (0.622-2.298) p=0.543

*OR values for top-bottom classification were based on controls only and were age adjusted.

**OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Supplemental Table VI-32: Joint association of high levels of androgens and gonadotrophins (top quantiles) with risk of breast cancer – cases that gave a sample more than 2 years before diagnosis.

Hormones		Joint association of androgens and gonadotrophins				
		Androstenedione		Testosterone	DHEAS	
		LH	FSH	LH	LH	FSH
		OR* (95% CI)	**OR* (95% CI)**	**OR* (95% CI)**	**OR* (95% CI)**	**OR* (95% CI)**
Not adjusted		2.127 (0.720-4.720) p=0.119	2.320 (0.863-6.889) p=0.108	2.248 (0.673-4.961) p=0.095	0.739 (0.211-1.885) p=0.592	1.350 (0.307-4.294) p=0.661
		OR (95% CI)**	**OR** (95% CI)**	**OR** (95% CI)**	**OR** (95% CI)**	**OR** (95% CI)**
Adjusted	Oestradiol	2.208 (0.802-5.605) p=0.105	2.334 (0.782-6.357) p=0.106	2.160 (0.786-5.464) p=0.114	0.744 (0.211-2.050) p=0.601	1.332 (0.289-4.592) p=0.673
	Oestrone	2.529 (0.906-6.559) p=0.062	2.523 (0.833-7.048) p=0.083	2.658 (0.947-6.977) p=0.051	0.813 (0.228-2.272) p=0.717	1.093 (0.242-3.614) p=0.894
	Androstenedione			1.796 (0.638-4.658) p=0.241	0.588 (0.164-1.650) p=0.355	1.095 (0.234-3.843) p=0.896
	Testosterone	1.694 (0.598-4.424) p=0.295	1.853 (0.607-5.153) p=0.250		0.518 (0.144-1.468) p=0.256	1.041 (0.222-3.676) p=0.953
	DHEAS	2.228 (0.795-5.780) p=0.108	2.346 (0.783-6.417) p=0.105	2.466 (0.855-6.667) p=0.081		
	Progesterone	2.092 (0.762-5.263) p=0.129	2.438 (0.813-6.688) p=0.091	2.218 (0.806-5.626) p=0.103	0.677 (0.192-1.866) p=0.491	1.276 (0.277-4.402) p=0.720
	SHBG	2.283 (0.822-5.858) p=0.094	2.311 (0.771-6.324) p=0.112	2.008 (0.677-5.456) p=0.186	0.712 (0.197-2.033) p=0.559	1.489 (0.316-5.364) p=0.568
	LH		3.002 (0.974-8.630) **p=0.044**			1.973 (0.409-7.413) p=0.342
	FSH	2.169 (0.789-5.479) p=0.112		2.288 (0.832-5.797) p=0.090	0.751 (0.212-2.070) p=0.612	
	ER-α SB	2.303 (0.832-5.875) p=0.090	2.508 (0.836-6.878) p=0.081	2.335 (0.846-5.938) p=0.083	0.743 (0.210-2.054) p=0.601	1.667 (0.357-5.909) p=0.460
	ER-β SB	2.165 (0.790-5.444) p=0.111	2.302 (0.770-6.272) p=0.112	2.188 (0.795-5.539) p=0.108	0.747 (0.211-2.058) p=0.606	1.190 (0.262-3.992) p=0.795
	AR SB	2.140 (0.782-5.370) p=0.116	2.426 (0.811-6.622) p=0.092	2.261 (0.824-5.715) p=0.094	0.736 0.209-2.022) p=0.587	1.216 (0.268-4.070) p=0.771

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Supplemental Table VI-33: Joint association of high levels of androgens and gonadotrophins (top quantiles) with risk of breast cancer – cases that gave a sample more than 2 years before diagnosis.

Hormones		Joint association of androgens and SHBG		
		Androstenedione	Testosterone	DHEAS
		SHBG		
		OR* (95% CI)	**OR* (95% CI)**	**OR* (95% CI)**
Not adjusted		0.276 (0.028-) p=0.217	0.703 (0.054-2.873) p=0.650	0.709 (0.277-3.091) p=0.659
		OR (95% CI)**	**OR** (95% CI)**	**OR** (95% CI)**
Adjusted	Oestradiol	not enough points	0.716 (0.108-2.781) p=0.670	0.787 (0.166-2.961) p=0.761
	Oestrone	not enough points	0.755 (0.115-2.930) p=0.719	0.766 (0.096-2.481) p=0.733
	Androstenedione		0.300 (0.016-1.609) p=0.256	0.635 (0.077-2.032) p=0.564
	Testosterone	not enough points		0.513 (0.093-2.632) p=0.400
	DHEAS	0.269 (0.008-1.373) p=0.209	0.666 (0.101-2.594) p=0.606	
	Progesterone			
	SHBG	0.251 (0.014-1.289) p=0.187	0.625 (0.094-2.439) p=0.551	0.604 (0.090-2.406) p=0.527
	LH	0.267 (0.015-1.357) p=0.205	0.703 (0.107-2.687) p=0.651	0.677 (0.103-2.609) p=0.618
	FSH	0.275 (0.015-1.396) p=0.215	0.703 (0.107-2.686) p=0.651	0.712 (0.108-2.734) p=0.663
	ER-α SB	0.275 (0.015-1.398) p=0.216	0.705 (0.108-2.698) p=0.654	0.710 (0.108-2.728) p=0.661
	ER-β SB	not enough points	0.654 (0.100-2.508) p=0.587	0.679 (0.103-2.613) p=0.621
	AR SB	not enough points	0.703 (0.107-2.686) p=0.651	0.708 (0.108-2.723) p=0.659

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Supplemental Table VI-34: Joint association of high levels of progesterone and gonadotrophins (top quantiles) with risk of breast cancer – cases that gave a sample more than 2 years before diagnosis.

Hormones		Joint association of progesterone and gonadotrophins	
		Progesterone	
		LH	FSH
		OR* (95% CI)	**OR** (95% CI)**
Not adjusted		0.403 (0.083-1.605) p=0.229	0.944 (0.179-4.049) p=0.942
		OR (95% CI)**	**OR** (95% CI)**
Adjusted	Oestradiol	0.393 (0.062-1.396) p=0.216	0.846 (0.127-3.344) p=0.832
	Oestrone	0.933 (0.261-2.631) p=0.904	1.279 (0.279-4.362) p=0.716
	Androstenedione	0.338 (0.053-1.220) p=0.154	0.675 (0.100-2.724) p=0.623
	Testosterone	0.286 (0.044-1.045) p=0.102	0.730 (0.107-3.012) p=0.697
	DHEAS	0.377 (0.058-1.421) p=0.210	0.949 (0.140-3.934) p=0.948
	Progesterone	0.374 (0.059-1.333) p=0.194	0.903 (0.135-3.642) p=0.898
	SHBG		
	LH		1.140 (0.169-4.738) p=0.871
	FSH	0.406 (0.064-1.442) p=0.232	
	ER-α SB	0.871 (0.245-2.427) p=0.809	1.708 (0.367-6.033) p=0.440
	ER-β SB	0.407 (0.064-1.449) p=0.234	0.959 (0.143-3.866) p=0.958
	AR SB	0.833 (0.235-2.309) p=0.748	1.339 (0.293-4.560) p=0.666

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Supplemental Table VI-35: Joint association of high levels of androgens and progesterone (top quantiles) with risk of breast cancer – cases that gave a sample more than 2 years before diagnosis.

Hormones		Joint association of gonadotrophins LH FSH
		OR (95% CI)**
Not adjusted		0.777 (0.317-1.567) p=0537
		OR (95% CI)**
Adjusted	Oestradiol	0.821 (0.343-1.751) p=0.631
	Oestrone	0.773 (0.324-1.645) p=0.530
	Androstenedione	0.815 (0.341-1.740) p=0.619
	Testosterone	0.873 (0.364-1.874) p=0.743
	DHEAS	0.777 (0.326-1.647) p=0.536
	SHBG	0.818 (0.342-1.744) p=0.624
	Progesterone	0.803 (0.336-1.711) p=0.593
	LH	
	FSH	
	ER-α SB	0.836 (0.349-1.783) p=0.662
	ER-β SB	0.808 (0.339-1.721) p=0.604
	AR SB	0.794 (0.333-1.690) p=0.574

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Supplemental Table VI-36: Joint association of high levels of androgens and progesterone (top quantiles) with risk of breast cancer – cases that gave a sample more than 2 years before diagnosis.

Hormones		Joint association of SHBG and progesterone SHBG Progesterone
		OR* (95% CI)
Not adjusted		0.539 (0.082-1.916) p=0.420
		OR (95% CI)**
Adjusted	Oestradiol	0.511 (0.079-1.895) p=0.384
	Oestrone	0.474 (0.074-1.716) p=0.327
	Androstenedione	0.480 (0.074-1.797) p=0.343
	Testosterone	0.421 (0.064-1.592) p=0.266
	DHEAS	0.511 (0.078-1.920) p=0.386
	Progesterone	
	SHBG	
	LH	0.514 (0.079-1.897) p=0.386
	FSH	0.533 (0.082-1.966) p=0.412
	ER-α SB	0.502 (0.078-1.827) p=0.366
	ER-β SB	0.487 (0.076-1.776) p=0.346
	AR SB	0.468 (0.073-1.687) p=0.318

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Supplemental Table VI-37: Joint association of high levels of sex hormone bidning globulin and gonadotrophins (top quantiles) with risk of breast cancer – cases that gave a sample more than 2 years before diagnosis.

Hormones		Joint association of SHBG and gonadotrophins	
		SHBG	
		LH	FSH
		OR* (95% CI)	**OR* (95% CI)**
Not adjusted		0.221 (0.013-1.121) p=0.145	0.565 (0.142-1.590) p=0.364
		OR (95% CI)**	**OR** (95% CI)**
Adjusted	Oestradiol	0.218 (0.012-1.089) p=0.142	0.777 (0.221-2.135) p=0.655
	Oestrone	0.232 (0.013-1.163) p=0.159	0.577 (0.133-1.749) p=0.386
	Androstenedione	not enough points	0.597 (0.138-1.801) p=0.415
	Testosterone	0.250 (0.014-1.273) p=0.184	0.440 (0.069-1.570) p=0.278
	DHEAS	0.223 (0.012-1.122) p=0.149	0.570 (0.057-1.713) p=0.374
	Progesterone		
	SHBG	0.226 (0.027-1.136) p=0.153	0.362 (0.140-1.268) p=0.175
	LH		0.611 (0.103-1.881) p=0.442
	FSH	0.200 (0.012-1.032) p=0.125	
	ER-α SB	0.236 (0.013-1.182) p=0.164	0.373 (0.059-1.309) p=0.189
	ER-β SB	0.209 (0.011-1.047) p=0.132	0.344 (0.054-1.211) p=0.156
	AR SB	0.221 (0.012-1.103) p=0.146	0.373 (0.059-1.310) p=0.189

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Appendix VII: Statistically non-significant joint associations of high levels of sex steroid receptor serum bioactivities with breast cancer risk

Supplemental Table VII-1: Joint association of high sex steroid receptor serum bioactivity (top quantiles) with risk of breast cancer – all cases.

Joint association of high SB of sex steroid receptors - all cases				
SB		**ER-α and ER-β SB**	**ER-α and AR SB**	**ER-β and AR SB**
		OR* (95% CI)	**OR* (95% CI)**	**OR* (95% CI)**
Not adjusted		1.399 (0.833-2.338) p=0.202	1.428 (0.833-2.338) p=0.199	1.713 (0.867-2.739) p=0.065
		OR (95% CI)**	**OR** (95% CI)**	**OR** (95% CI)**
Adjusted	Oestradiol	1.421 (0.840-2.377) p=0.184	1.411 (0.840-2.423) p=0.216	1.698 (0.952-3.030) p=0.070
	Oestrone	1.323 (0.781-2.213) p=0.291	1.396 (0.781-2.407) p=0.234	1.597 (0.891-2.839) p=0.111
	Androstenedione	1.231 (0.745-2.144) p=0.370	1.416 (0.745-2.449) p=0.217	1.586 (0.882-2.827) p=0.119
	Testosterone	1.360 (0.799-2.286) p=0.250	1.324 (0.799-2.290) p=0.320	1.622 (0.755-2.290) p=0.106
	DHEAS	1.399 (0.767-2.177) p=0.201	1.427 (0.829-2.446) p=0.200	1.715 (0.924-2.937) p=0.066
	SHBG	1.300 (0.767-2.177) p=0.322	1.415 (0.767-2.440) p=0.214	1.624 (0.906-2.888) p=0.099
	Progesterone	1.386 (0.821-2.313) p=0.215	1.397 (0.804-2.395) p=0.228	1.743 (0.973-3.100) p=0.059
	LH	1.388 (0.821-2.317) p=0.214	1.388 (0.798-2.383) p=0.238	1.676 (0.939-2.970) p=0.077
	FSH	1.393 (0.824-2.325) p=0.208	1.420 (0.816-2.439) p=0.207	1.708 (0.958-3.023) p=0.066
	ER-α SB			1.420 (0.744-2.693) p=0.284
	ER-β SB		1.516 (0.905-2.795) p=0.183	
	AR SB	1.345 (0.764-2.346) p=0.298		

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Supplemental Table VII-2: Joint association of high sex steroid receptor serum bioactivity (top quantiles) with risk of breast cancer - cases that gave a sample less than 2 years before diagnosis.

Joint association of high SB of sex steroid receptors – less than 2 years before diagnosis				
SB		**ER-α and ER-β SB**	**ER-α and AR SB**	**ER-β and AR SB**
		OR* (95% CI)	**OR* (95% CI)**	**OR* (95% CI)**
Not adjusted		0.931 (0.454-1.877) p=0.847	0.956 (0.436-2.057) p=0.908	1.060 (0.407-2.104) p=0.888
		OR (95% CI)**	**OR** (95% CI)**	**OR** (95% CI)**
Adjusted	Oestradiol	0.941 (0.430-1.888) p=0.871	0.962 0.421-1.995) p=0.921	1.048 (0.434-2.273) p=0.910
	Oestrone	0.869 (0.396-1.748) p=0.708	0.922 (0.401-1.921) p=0.836	0.954 (0.391-2.087) p=0.910
	Androstenedione	0.792 (0.347-1.635) p=0.551	0.964 (0.419-2.013) p=0.925	1.003 (0.412-2.196) p=0.994
	Testosterone	0.891 (0.405-1.799) p=0.759	0.863 (0.374-1.807) p=0.710	0.991 (0.405-2.186) p=0.984
	DHEAS	0.932 (0.427-1.866) p=0.851	0.951 (0.417-1.967) p=0.898	1.046 (0.432-2.274) p=0.915
	SHBG	0.793 (0.362-1.597) p=0.537	0.882 (0.38-1.834) p=0.751	0.901 (0.372-1.961) p=0.802
	Progesterone	0.920 (0.421-1.841) p=0.822	0.932 (0.408-1.929) p=0.857	1.076 (0.444-2.342) p=0.862
	LH	0.931 (0.426-1.846) p=0.848	0.946 (0.415-1.959) p=0.888	1.049 (0.434-2.275) p=0.909
	FSH	0.928 (0.425-1.858) p=0.840	0.951 (0.417-1.971) p=0.899	1.060 (0.439-2.298) p=0.889
	ER-α SB			1.014 (0.388-2.439) p=0.976
	ER-β SB		1.243 (0.509-2.829) p=0.615	
	AR SB	0.890 (0.392-1.869) p=0.768		

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Appendix VIII: Statistically non-significant joint associations of high levels of sex steroid receptor serum bioactivity and hormones with breast cancer risk

Supplemental Table VIII-1: Joint association of high ER-α serum bioactivity and oestrogens (top quantiles) with risk of breast cancer - all cases.

SB - Hormones		Joint association of SB and oestrogens	
		ER-α SB	
		Oestradiol	Oestrone
		OR* (95% CI)	**OR* (95% CI)**
Not adjusted		1.412 (0.672-3.144) p=0.375	1.573 (0.872-3.632) p=0.333
		OR (95% CI)**	**OR** (95% CI)**
Adjusted	Oestradiol		1.634 (0.775-3.394) p=0.341
	Oestrone	1.540 (0.691-3.365) p=0.280	
	Androstenedione	1.531 (0.682-3.366) p=0.291	1.738 (0.819-3.634) p=0.325
	Testosterone	1.194 (0.529-2.619) p=0.661	1.664 (0.779-3.501) p=0.552
	DHEAS	1.414 (0.645-3.007) p=0.373	1.581 (0.754-3.250) p=0.332
	SHBG	1.463 (0.662-3.144) p=0.333	1.579 (0.748-3.274) p=0.439
	Progesterone	1.475 (0.668-3.176) p=0.323	1.658 (0.758-3.449) p=0.287
	LH	1.360 (0.619-2.899) p=0.431	1.590 (0.757-3.276) p=0.420
	FSH	1.428 (0.650-3.047) p=0.361	1.583 (0.755-3.257) p=0.315
	ER-α SB		
	ER-β SB	1.404 (0.637-3.005) p=0.387	1.615 (0.744-3.451) p=0.217
	AR SB	1.374 (0.612-2.996) p=0.428	1.533 (0.716-3.223) p=0.262

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Supplemental Table VIII-2: Joint association of high ER-α serum bioactivity and androgens (top quantiles) with risk of breast cancer - all cases.

SB/Hormones		Joint association of SB and androgens	
		ER-α SB	
		Androstenedione	DHEAS
		OR* (95% CI)	**OR* (95% CI)**
Not adjusted		1.882 (1.09-4.638) p=0.077	0.976 (0.425-2.421) p=0.956
		OR (95% CI)**	**OR** (95% CI)**
Adjusted	Oestradiol	2.083 (1.007-4.312) **p=0.046**	0.946 (0.375-2.204) p=0.901
	Oestrone	1.872 (0.918-3.787) p=0.080	0.956 (0.378-2.230) p=0.919
	Androstenedione		0.817 (0.318-1.938) p=0.658
	Testosterone	1.391 (0.652-2.945) p=0.388	0.706 (0.270-1.711) p=0.454
	DHEAS	1.877 (0.917-3.811) p=0.081	
	SHBG	1.785 (0.873-3.620) p=0.108	0.943 (0.372-2.209) p=0.895
	Progesterone	1.950 (0.944-4.017) p=0.068	0.936 (0.361-2.254) p=0.885
	LH	1.866 (0.914-3.783) p=0.083	0.989 (0.391-2.311) p=0.981
	FSH	1.868 (0.915-3.787) p=0.082	0.961 (0.380-2.247) p=0.929
	ER-α SB		
	ER-β SB	1.969 (0.963-4.119) p=0.071	0.960 (0.375-2.274) p=0.929
	AR SB	1.834 (0.884-3.777) p=0.099	0.911 (0.353-2.169) p=0.838

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Supplemental Table VIII-3: Joint association of high ER-α serum bioactivity and progesterone (top quantiles) with risk of breast cancer - all cases.

SB - Hormones		Joint association of SB and progesterone ER-α SB Progesterone
		OR* (95% CI)
Not adjusted		1.443 (0.613-3.558) p=0.415
		OR (95% CI)**
Adjusted	Oestradiol	1.523 (0.603-3.723) p=0.359
	Oestrone	1.377 (0.551-3.305) p=0.478
	Androstenedione	1.121 (0.438-2.756) p=0.805
	Testosterone	0.963 (0.374-2.380) p=0.936
	DHEAS	1.436 (0.559-3.556) p=0.437
	SHBG	1.520 (0.605-3.673) p=0.357
	Progesterone	
	LH	1.435 (0.574-3.455) p=0.423
	FSH	1.412 (0.564-3.401) p=0.445
	ER-α SB	
	ER-β SB	1.442 (0.568-3.518) p=0.426
	AR SB	1.363 (0.528-3.382) p=0.509

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Supplemental Table VIII-4: Joint association of high ER-α serum bioactivity and sex hormone-binding globulin (top quantiles) with risk of breast cancer - all cases.

SB/Hormones		Joint association of SB and SHBG ER-α SB SHBG
		OR* (95% CI)
Not adjusted		1.524 (0.999-4.658) p=0.443
		OR (95% CI)**
Adjusted	Oestradiol	1.759 (0.604-5.009) p=0.286
	Oestrone	1.430 (0.462-4.203) p=0.516
	Androstenedione	1.577 (0.506-4.671) p=0.412
	Testosterone	1.421 (0.452-4.244) p=0.530
	DHEAS	1.526 (0.495-4.464) p=0.441
	SHBG	
	Progesterone	1.747 (0.601-4.966) p=0.291
	LH	1.580 (0.510-4.651) p=0.407
	FSH	1.525 (0.495-4.461) p=0.442
	ER-α SB	
	ER-β SB	1.515 (0.490-4.440) p=0.450
	AR SB	1.453 (0.464-4.323) p=0.503

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Supplemental Table VIII-5: Joint association of ER-α serum bioactivity and gonadotrophins (top quantiles) with risk of breast cancer - all cases.

SB - Hormones		Joint association of SB and gonadotrophins	
		ER-α SB	
		LH	FSH
		OR* (95% CI)	**OR* (95% CI)**
Not adjusted		0.669 (0.257-1.74) p=0.405	1.168 (0.465-3.159) p=0.748
		OR (95% CI)**	**OR** (95% CI)**
Adjusted	Oestradiol	0.540 (0.175-1.391) p=0.234	0.983 (0.337-2.577) 0.973
	Oestrone	0.524 (0.170-1.353) p=0.213	0.954 (0.326-2.509) p=0.926
	Androstenedione	0.556 (0.179-1.438) p=0.258	1.084 (0.370-2.862) p=0.874
	Testosterone	0.550 (0.176-1.433) p=0.253	1.124 (0.381-2.988) p=0.820
	DHEAS	0.668 (0.237-1.637) p=0.404	1.183 (0.432-3.005) p=0.730
	SHBG	0.664 (0.235-1.640) p=0.401	1.162 (0.511-2.973) p=0.759
	Progesterone	0.547 (0.177-1.408) p=0.244	1.020 (0.349-2.689) p=0.969
	LH		1.433 (0.511-3.761) p=0.473
	FSH	0.678 (0.239-1.680) p=0.427	
	ER-α SB		
	ER-β SB	0.540 (0.181-1.401) p=0.237	0.978 (0.587-2.589) p=0.966
	AR SB	0.540 (0.190-1.392) p=0.235	0.976 (0.414-2.563) p=0.962

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Supplemental Table VIII-6: Joint association of high ER-β serum and oestrogens (top quantiles) with risk of breast cancer - all cases.

SB - Hormones		Joint association of SB and oestrogens	
		ER-β SB	
		Oestradiol	Oestrone
		OR* (95% CI)	**OR* (95% CI)**
Not adjusted		0.752 (0.445-2.056) p=0.530	1.154 (0.610-2.571) p=0.648
		OR (95% CI)**	**OR** (95% CI)**
Adjusted	Oestradiol		1.197 (0.571-2.419) p=0.573
	Oestrone	0.728 (0.278-1.708) p=0.486	
	Androstenedione	0.729 (0.275-1.742) p=0.496	1.210 (0.574-2.461) p=0.541
	Testosterone	0.608 (0.228-1.453) p=0.284	1.174 (0.554-2.402) p=0.350
	DHEAS	0.749 (0.286-1.754) p=0.524	1.160 p=0.644
	SHBG	0.702 (0.268-1.652) p=0.439	1.057 (0.505-2.131) p=0.486
	Progesterone	0.783 (0.298-1.853) p=0.594	1.195 (0.570-2.416) p=0.710
	LH	0.742 (0.283-1.741) p=0.512	1.124 (0.538-2.259) p=0.612
	FSH	0.757 (0.289-1.776) p=0.541	1.169 (0.560-2.351) p=0.663
	ER-α SB	0.684 (0.259-1.731) p=0.410	1.017 (0.476-2.093) p=0.965
	ER-β SB		
	AR SB	0.763 (0.287-1.824) p=0.561	1.112 (0.524-2.270) p=0.774

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Supplemental Table VIII-7: Joint association of high ER-β serum bioactivity and androgens (top quantiles) with risk of breast cancer - all cases.

SB - Hormones		Joint association of SB and androgens		
		ER-β SB		
		Androstenedione	Testosterone	DHEAS
		OR* (95% CI)	**OR* (95% CI)**	**OR* (95% CI)**
Not adjusted		1.604 (0.760-3.395) p=0.200	1.856 (1.053-4.377) p=0.084	1.108 (0.438-2.574) p=0.820
		OR (95% CI)**	**OR** (95% CI)**	**OR** (95% CI)**
Adjusted	Oestradiol	1.665 (0.789-3.461) p=0.172	1.932 (0.940-3.995) p=0.070	1.263 (0.484-3.110) p=0.617
	Oestrone	1.504 (0.699-3.165) p=0.285	1.753 (0.842-3.616) p=0.128	1.159 (0.450-2.800) p=0.748
	Androstenedione		1.429 (0.664-) p=0.354	0.927 (0.351-2.290) p=0.872
	Testosterone	1.044 (0.467-2.278) p=0.914		0.820 (0.310-2.034) p=0.675
	DHEAS	1.597 (0.756-3.316) p=0.211	1.859 (0.906-3.787) p=0.087	
	SHBG	1.442 (0.686-2.974) p=0.324	1.672 (0.819-3.386) p=0.153	1.044 (0.408-2.491) p=0.925
	Progesterone	1.521 (0.702-3.225) p=0.276	1.829 (0.891-3.727) p=0.096	1.144 (0.430-2.869) p=0.778
	LH	1.557 (0.742-3.207) p=0.232	1.863 (0.914-3.768) p=0.083	1.097 (0.430-2.611) p=0.838
	FSH	1.596 (0.761-3.281) p=0.206	1.851 (0.910-3.736) p=0.085	1.099 (0.431-2.611) p=0.835
	ER-α SB	1.432 (0.669-3.035) p=0.351	1.693 (0.816-3.483) p=0.152	1.061 (0.405-2.606) p=0.900
	ER-β SB			
	AR SB	1.649 (0.764-3.509) p=0.195	1.953 (0.933-4.073) p=0.073	1.025 (0.393-2.500) p=0.958

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Supplemental Table VIII-8: Joint association of high ER-β serum bioctivity and progesterone (top quantiles) with risk of breast cancer - all cases.

SB - Hormones		Joint association of SB and progesterone ER-β SB Progesterone
		OR* (95% CI)
Not adjusted		0.751 (0.287-1.973) p=0.562
		OR (95% CI)**
Adjusted	Oestradiol	0.838 (0.289-2.159) p=0.726
	Oestrone	0.882 (0.327-2.169) p=0.792
	Androstenedione	0.658 (0.222-1.739) p=0.418
	Testosterone	0.521 (0.176-1.367) p=0.205
	DHEAS	0.702 (0.239-1.828) p=0.488
	SHBG	0.733 (0.256-1.851) p=0.531
	Progesterone	
	LH	0.728 (0.254-1.833) p=0.522
	FSH	0.740 (0.259-1.862) p=0.543
	ER-α SB	0.792 (0.293-1.943) p=0.623
	ER-β SB	
	AR SB	0.885 (0.324-2.212) p=0.801

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Supplemental Table VIII-9: Joint association of high ER-β serum bioactivity and gonadotrophins (top quantiles) with risk of breast cancer - all cases.

SB - Hormones		Joint association of SB and gonadotrophins	
		ER-β SB	
		LH	FSH
		OR* (95% CI)	**OR* (95% CI)**
Not adjusted		0.615 (0.207-1.612) p=0.350	1.637 (0.621-3.813) p=0.282
		OR (95% CI)**	**OR** (95% CI)**
Adjusted	Oestradiol	0.605 (0.195-1.575) p=0.335	1.631 (0.646-4.017) p=0.287
	Oestrone	0.609 (0.196-1.589) p=0.342	1.654 (0.654-4.082) p=0.274
	Androstenedione	0.641 (0.206-1.677) p=0.395	1.695 (0.668-4.195) p=0.254
	Testosterone	0.570 (0.182-1.505) p=0.287	1.588 (0.621-3.956) p=0.321
	DHEAS	0.616 (0.198-1.599) p=0.350	1.645 (0.652-4.048) p=0.278
	SHBG	0.559 (0.180-1.461) p=0.266	1.445 (0.413-3.565) p=0.423
	Progesterone	0.607 (0.196-1.579) p=0.338	1.629 (0.645-4.012) p=0.288
	LH		1.828 (0.717-4.562) p=0.195
	FSH	0.623 (0.200-1.626) p=0.365	
	ER-α SB	0.551 (0.195-1.454) p=0.258	1.528 (0.334-3.785) p=0.359
	ER-β SB		
	AR SB	0.606 (0.202-1.577) p=0.336	1.600 (0.371-3.953) p=0.308

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Supplemental Table VIII-10: Joint association of high ER-β serum bioactivity and sex hormone-binding globulin (top quantiles) with risk of breast cancer - all cases.

SB/Hormones		Joint association of SB and SHBG SHBG ER-β SB
		OR* (95% CI)
Not adjusted		3.428 (0.840-17.256) p=0.094
		OR (95% CI)**
Adjusted	Oestradiol	4.026 (1.049-19.255) p=0.051
	Oestrone	3.361 (0.811-16.610) p=0.101
	Androstenedione	3.828 (0.917-19.033) p=0.071
	Testosterone	3.973 (0.946-19.815) p=0.064
	DHEAS	3.427 (0.830-16.880) p=0.095
	SHBG	
	Progesterone	3.353 (0.811-16.529) p=0.101
	LH	3.331 (0.805-16.439) p=0.103
	FSH	3.382 (0.817-16.699) p=0.099
	ER-α SB	3.103 (0.742-15.427) p=0.128
	ER-β SB	
	AR SB	3.276 (0.784-16.287) p=0.110

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Supplemental Table VIII-11: Joint association of high AR serum bioactivity and oestrogens (top quantiles) with risk of breast cancer - all cases.

SB/Hormones		Joint assocation of SB and oestrogens
		AR SB
		Oestradiol
		OR* (95% CI)
Not adjusted		1.537 p=0.257
		OR (95% CI)**
Adjusted	Oestradiol	
	Oestrone	1.692 (0.775-3.647) p=0.179
	Androstenedione	1.589 (0.614-3.398) p=0.233
	Testosterone	1.336 (0.699-2.843) p=0.455
	DHEAS	1.537 (0.724-3.226) p=0.258
	SHBG	1.468 (0.681-3.096) p=0.315
	Progesterone	1.502 (0.717-3.157) p=0.286
	LH	1.507 (0.668-3.166) p=0.282
	FSH	1.557 (0.708-3.276) p=0.246
	ER-α SB	1.474 (0.667-3.196) p=0.327
	ER-β SB	1.527 (0.708-3.220) p=0.269
	AR SB	

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Supplemental Table VIII-12: Joint association of high AR serum bioactivity and androgens (top quantiles) with risk of breast cancer - all cases.

SB/Hormones		Joint association of SB and androgens
		AR SB
		DHEAS
		OR* (95% CI)
Not adjusted		1.148 (0.487-2.641) p=0.747
		OR (95% CI)**
Adjusted	Oestradiol	1.206 (0.411-2.788) p=0.667
	Oestrone	1.027 (0.427-2.317) p=0.951
	Androstenedione	1.021 (0.344-2.399) p=0.962
	Testosterone	0.863 (0.429-2.049) p=0.743
	DHEAS	
	SHBG	1.072 (0.442-2.451) p=0.872
	Progesterone	1.064 (0.450-2.493) p=0.889
	LH	1.120 (0.442-2.554) p=0.791
	FSH	1.132 (0.460-2.584) p=0.773
	ER-α SB	1.097 (0.441-2.577) p=0.836
	ER-β SB	1.131 (0.460-2.626) p=0.779
	AR SB	

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Supplemental Table VIII-13: Joint association of high AR serum bioactivity and androgens (top quantiles) with risk of breast cancer - all cases.

SB/Hormones		Joint association of SB and SHBG AR SB SHBG
		OR* (95% CI)
Not adjusted		1.187 (0.456-3.798) p=0.744
		OR (95% CI)**
Adjusted	Oestradiol	1.314 (0.435-3.700) p=0.610
	Oestrone	1.272 (0.420-3.594) p=0.653
	Androstenedione	1.595 (0.513-4.703) p=0.398
	Testosterone	1.298 (0.423-3.708) p=0.631
	DHEAS	1.187 (0.398-3.251) p=0.743
	SHBG	
	Progesterone	1.164 (0.390-3.186) p=0.773
	LH	1.196 (0.400-3.283) p=0.734
	FSH	1.186 (0.398-3.247) p=0.745
	ER-α SB	1.184 (0.386-3.379) p=0.756
	ER-β SB	1.307 (0.431-3.687) p=0.618
	AR SB	

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Supplemental Table VIII-14: Joint association of high AR serum bioactivity and progesterone (top quantiles) with risk of breast cancer - all cases.

SB - Hormones		Joint association of SB and progesterone AR SB Progesterone
		OR* (95% CI)
Not adjusted		1.739 (0.239-3.907) p=0.190
		OR (95% CI)**
Adjusted	Oestradiol	1.554 (0.602-3.651) p=0.314
	Oestrone	1.647 (0.704-3.788) p=0.239
	Androstenedione	1.433 (0.443-3.350) p=0.406
	Testosterone	1.097 (0.683-2.637) p=0.838
	DHEAS	1.823 (0.734-4.342) p=0.174
	SHBG	1.765 (0.751-4.088) p=0.184
	Progesterone	
	LH	1.706 (0.700-3.918) p=0.207
	FSH	1.718 (0.736-3.960) p=0.203
	ER-α SB	1.694 (0.700-4.059) p=0.234
	ER-β SB	1.740 (0.736-4.060) p=0.198
	AR SB	

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Supplemental Table VIII-15: Joint association of high AR serum bioactivity and gonadotrophins (top quantiles) with risk of breast cancer - all cases.

SB - Hormones		Joint association of SB and gonadotrophins	
		AR SB	
		LH	FSH
		OR* (95% CI)	**OR* (95% CI)**
Not adjusted		0.594 (0.153-2.096) p=0.434	0.892 (0.343-2.499) p=0.821
		OR (95% CI)**	**OR** (95% CI)**
Adjusted	Oestradiol	0.583 (0.133-1.937) p=0.418	0.894 (0.341-2.308) p=0.824
	Oestrone	0.597 (0.132-1.989) p=0.439	0.899 (0.310-2.326) p=0.833
	Androstenedione	0.601 (0.123-2.011) p=0.447	0.991 (0.339-2.574) p=0.986
	Testosterone	0.562 (0.129-1.899) p=0.392	0.988 (0.311-2.577) p=0.981
	DHEAS	0.594 (0.134-1.971) p=0.433	0.893 (0.307-2.306) p=0.822
	SHBG	0.618 (0.136-2.085) p=0.473	0.884 (0.572-2.306) p=0.808
	Progesterone	0.580 (0.132-1.924) p=0.412	0.903 (0.308-2.338) p=0.839
	LH		0.980 (0.295-2.557) p=0.968
	FSH	0.607 (0.129-2.036) p=0.456	
	ER-α SB	0.490 (0.110-1.609) p=0.282	0.856 (0.632-2.215) p=0.758
	ER-β SB	0.583 (0.130-1.946) p=0.419	0.878 (0.660-2.277) p=0.656
	AR SB		

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Supplemental Table VIII-16: Joint association of high ER-α serum bioactivity and oestrogens (top quantiles) with risk of breast cancer – cases that gave a sample less than 2 years before diagnosis.

SB - Hormones		Joint association of SB and oestrogens	
		ER-α SB	
		Oestradiol	Oestrone
		OR* (95% CI)	**OR* (95% CI)**
Not adjusted		0.895 (0.275-2.600) p=0.845	1.293 (0.643-3.995) p=0.597
		OR (95% CI)**	**OR** (95% CI)**
Adjusted	Oestradiol		1.368 (0.481-3.407) p=0.523
	Oestrone	0.946 (0.263-2.697) p=0.923	
	Androstenedione	0.982 (0.273-2.808) p=0.975	1.455 (0.509-3.653) p=0.448
	Testosterone	0.780 (0.216-2.233) p=0.669	1.319 (0.458-3.335) p=0.578
	DHEAS	0.898 (0.254-2.497) p=0.850	1.301 (0.460-3.211) p=0.589
	SHBG	0.941 (0.264-2.656) p=0.916	1.239 (0.435-3.085) p=0.663
	Progesterone	0.925 (0.260-2.597) p=0.891	1.353 (0.476-3.374) p=0.538
	LH	0.881 (0.249-2.453) p=0.824	1.306 (0.462-3.227) p=0.584
	FSH	0.911 (0.257-2.538) p=0.869	1.307 (0.462-3.228) p=0.583
	ER-α SB		
	ER-β SB	0.988 (0.277-2.791) p=0.983	1.626 (0.554-4.253) p=0.342
	AR SB	0.883 (0.244-535) p=0.831	1.306 (0.453-3.302) p=0.592

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Supplemental Table VIII-17: Joint association of high ER-α serum bioactivity and androgens (top quantiles) with risk of breast cancer – cases that gave a sample less than 2 years before diagnosis.

SB - Hormones		Joint association of SB and androgens		
		ER-α SB		
		Androstenedione	Testosterone	DHEAS
		OR* (95% CI)	**OR* (95% CI)**	**OR* (95% CI)**
Not adjusted		1.617 (0.706-4.505) p=0.300	1.704 (10.806-4.663) p=0.226	0.686 (0.202-2.506) p=0.558
		OR (95% CI)**	**OR** (95% CI)**	**OR** (95% CI)**
Adjusted	Oestradiol	1.825 (0.680-4.447) p=0.202	1.829 (0.725-4.259) p=0.175	0.674 (0.153-2.100) p=0.541
	Oestrone	1.627 (0.612-3.901) p=0.296	1.645 (0.655-3.794) p=0.260	0.661 (0.150-2.071) p=0.523
	Androstenedione		1.383 (0.532-3.312) p=0.482	0.582 (0.130-1.858) p=0.409
	Testosterone	1.207 (0.434-3.054) p=0.702		0.498 (0.110-1.619) p=0.294
	DHEAS	1.588 (0.595-3.820) p=0.322	1.668 (0.660-3.879) p=0.251	
	SHBG	1.491 (0.560-3.583) p=0.392	1.679 (0.666-3.896) p=0.244	0.636 (0.144-1.994) p=0.485
	Progesterone	1.640 (0.608-4.009) p=0.297	1.628 (0.644-3.783) p=0.274	0.656 (0.146-2.112) p=0.522
	LH	1.608 (0.606-3.846) p=0.306	1.707 (0.682-3.919) p=0.224	0.701 (0.159-2.186) p=0.582
	FSH	1.608 (0.603-3.874) p=0.309	1.705 (0.679-3.937) p=0.228	0.679 (0.154-2.122) p=0.549
	ER-α SB			
	ER-β SB	2.164 (0.778-5.574) p=0.119	2.034 (0.792-4.858) p=0.120	0.780 (0.176-2.482) p=0.704
	AR SB	1.617 (0.599-3.950) p=0.310	1.742 (0.680-4.126) p=0.222	0.673 (0.151-2.146) p=0.545

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Supplemental Table VIII-18: Joint association of high ER-α serum bioactivity and sex hormone-binding globulin (top quantiles) with risk of breast cancer – cases that gave a sample less than 2 years before diagnosis.

SB/Hormones		Joint association of SB and SHBG ER-α SB SHBG
		OR* (95% CI)
Not adjusted		0.484 (0.062-2.710) p=0.496
		OR (95% CI)**
Adjusted	Oestradiol	0.505 (0.027-2.854) p=0.524
	Oestrone	0.441 (0.023-2.483) p=0.445
	Androstenedione	0.495 (0.026-2.799) p=0.513
	Testosterone	0.400 (0.021-2.290) p=0.395
	DHEAS	0.485 (0.062-2.702) p=0.498
	SHBG	
	Progesterone	0.458 (0.024-2.562) p=0.465
	LH	0.507 (0.027-2.844) p=0.526
	FSH	0.486 (0.024-2.708) p=0.499
	ER-α SB	
	ER-β SB	0.449 (0.024-2.505) p=0.454
	AR SB	0.479 (0.026-2.687) p=0.492

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Supplemental Table VIII-19: Joint association of high ER-α serum bioactivity and progesterone (top quantiles) with risk of breast cancer – cases that gave a sample less than 2 years before diagnosis.

SB - Hormones		Joint association of SB and progesterone
		ER-α SB
		Progesterone
		OR* (95% CI)
Not adjusted		1.219 (0.373-3.765) p=0.738
		OR (95% CI)**
Adjusted	Oestradiol	1.357 (0.366-4.106) p=0.611
	Oestrone	1.114 (0.301-3.332) p=0.857
	Androstenedione	0.943 (0.250-2.910) p=0.924
	Testosterone	0.813 (0.215-2.508) p=0.735
	DHEAS	1.164 (0.309-3.590) p=0.804
	SHBG	1.283 (0.347-3.863) p=0.677
	Progesterone	
	LH	1.228 (0.335-3.644) p=0.729
	FSH	1.192 (0.324-3.557) p=0.768
	ER-α SB	
	ER-β SB	1.479 (0.394-4.572) p=0.521
	AR SB	1.229 (0.3.24-3.831) p=0.738

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Supplemental Table VIII-20: Joint association of high ER-α serum bioactivity and progesterone (top quantiles) with risk of breast cancer – cases that gave a sample less than 2 years before diagnosis.

SB - Hormones		Joint association of SB and gonadotrophins	
		ER-α SB	
		LH	FSH
		OR* (95% CI)	**OR* (95% CI)**
Not adjusted		0.872 (0.268-2.502) p=0.811	0.972 (0262-3.492) p=0.965
		OR (95% CI)**	**OR** (95% CI)**
Adjusted	Oestradiol	0.626 (0.143-1.927) p=0.465	0.625 (0.096-2.347) p=0.543
	Oestrone	0.604 (0.138-1.870) p=0.433	0.599 (0.092-2.268) p=0.509
	Androstenedione	0.672 (0.153-2.083) p=0.536	0.763 (0.117-2.895) p=0.728
	Testosterone	0.625 (0.142-1.948) p=0.466	0.737 (0.113-2.809) p=0.695
	DHEAS	0.869 (0.992-2.433) p=0.806	0.992 (0.222-3.220) p=0.990
	SHBG	0.870 (0.243-2.458) p=0.809	0.989 (0.220-3.239) p=0.985
	Progesterone	0.636 (0.654-1.956) p=0.479	0.654 (0.100-2.472) p=0.584
	LH		1.093 (0.240-3.677) p=0.895
	FSH	0.880 (0.246-2.493) p=0.825	
	ER-α SB		
	ER-β SB	0.724 (0.164-2.271) p=0.618	0.723 (0.110-2.766) p=0.677
	AR SB	0.639 (0.146-1.973) p=0.486	0.636 (0.098-2.349) p=0.559

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Supplemental Table VIII-21: Joint association of high ER-β serum bioactivity and oestrogen (top quantiles) with risk of breast cancer – cases that gave a sample less than 2 years before diagnosis.

SB - Hormones		Joint association of SB and oestrogens	
		ER-β SB	
		Oestradiol	Oestrone
		OR* (95% CI)	**OR* (95% CI)**
Not adjusted		0.205 (0.027-1.089) p=0.125	0.854 (0.333-2.516) p=0.756
		OR (95% CI)**	**OR** (95% CI)**
Adjusted	Oestradiol		0.878 (0.287-2.221) p=0.797
	Oestrone	0.200 (0.009-0.991) p=0.120	
	Androstenedione	0.206 (0.011-1.033) p=0.129	0.878 (0.285-2.240) p=0.800
	Testosterone	0.177 (0.010-0.888) p=0.096	0.831 (0.269-2.137) p=0.722
	DHEAS	0.204 (0.011-1.006) p=0.124	0.859 (0.282-2.163) p=0.765
	SHBG	0.190 (0.011-0.945) p=0.109	0.726 (0.238-1.835) p=0.531
	Progesterone	0.209 (0.012-1.040) p=0.131	0.877 (0.287-2.222) p=0.798
	LH	0.202 (0.011-0.998) p=0.122	0.845 (0.277-2.127) p=0.74
	FSH	0.208 (0.011-1.026) p=0.128	0.871 (0.285-2.198) p=0.787
	ER-α SB	0.202 (0.011-1.009) p=0.123	0.860 (0.277-2.221) p=0.771
	ER-β SB		
	AR SB	0.213 (0.012-1.065) p=0.136	0.847 (0.275-2.167) p=0.748

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Supplemental Table VIII-22: Joint association of high ER-β serum bioactivity and androgens (top quantiles) with risk of breast cancer – cases that gave a sample less than 2 years before

SB - Hormones		Joint association of SB and androgens		
		ER-β SB		
		Androstenedione	Testosterone	DHEAS
		OR* (95% CI)	**OR* (95% CI)**	**OR* (95% CI)**
Not adjusted		1.118 (0.402-3.150) p=0.830	1.088 (0.438-3.462) p=0.871	0.514 (0.111-2.235) p=0.384
		OR (95% CI)**	**OR** (95% CI)**	**OR** (95% CI)**
Adjusted	Oestradiol	1.180 (0.379-3.086) p=0.752	1.161 (0.372-3.042) p=0.776	0.603 (0.092-2.286) p=0.515
	Oestrone	1.071 (0.341-2.820) p=0.897	1.044 (0.333-2.750) p=0.936	0.511 (0.078-1.914) p=0.385
	Androstenedione		0.887 (0.275-2.429) p=0.827	0.426 (0.064-1.638) p=0.277
	Testosterone	0.807 (0.250-2.203) p=0.694		0.380 (0.057-1.459) p=0.217
	DHEAS	1.084 (0.347-2.843) p=0.878	1.058 (0.339-2.772) p=0.915	
	SHBG	0.945 (0.305-2.456) p=0.914	0.927 (0.298-2.410) p=0.884	0.460 (0.071-1.697) p=0.312
	Progesterone	1.105 (0.351-2.926) p=0.851	1.014 (0.323-2.669) p=0.979	0.470 (0.071-1.839) p=0.340
	LH	1.084 (0.349-2.813) p=0.877	1.088 (0.351-2.821) p=0.871	0.511 (0.079-1.875) p=0.380
	FSH	1.111 (0.358-2.883) p=0.840	1.085 (0.350-2.814) p=0.875	0.510 (0.079-1.874) p=0.380
	ER-α SB	1.166 (0.363-3.184) p=0.777	1.106 (0.349-2.963) p=0.851	0.553 (0.084-2.108) p=0.448
	ER-β SB			
	AR SB	1.170 (0.370-3.130) p=0.768	1.151 (0.363-3.082) p=0.793	0.490 (0.075-1.862) p=0.361

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Supplemental Table VIII-23: Joint association of high ER-β serum bioactivity and progesterone (top quantiles) with risk of breast cancer – cases that gave a sample less than 2 years before diagnosis.

SB - Hormones		Joint association of SB and progesterone ER-β SB Progesterone
		OR* (95% CI)
Not adjusted		0.471 (0.104-2.067) p=0.324
		OR (95% CI)**
Adjusted	Oestradiol	0.551 (0.085-2.058) p=0.441
	Oestrone	0.436 (0.067-1.623) p=0.282
	Androstenedione	0.417 (0.063-1.604) p=0.265
	Testosterone	0.338 (0.051-1.288) p=0.165
	DHEAS	0.415 (0.063-1.573) p=0.259
	SHBG	0.440 (0.068-1.616) p=0.284
	Progesterone	
	LH	0.465 (0.072-1.699) p=0.317
	FSH	0.464 (0.072-1.700) p=0.316
	ER-α SB	0.470 (0.072-1.746) p=0.327
	ER-β SB	
	AR SB	0.492 (0.075-1.849) p=0.361

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Supplemental Table VIII-24: Joint association of high ER-β serum bioactivity and progesterone (top quantiles) with risk of breast cancer – cases that gave a sample less than 2 years before diagnosis.

SB/Hormones		Joint association of SB and SHBG ER-β SB SHBG
		OR* (95% CI)
Not adjusted		not enough points
		OR (95% CI)**
Adjusted	Oestradiol	not enough points
	Oestrone	
	Androstenedione	
	Testosterone	
	DHEAS	
	SHBG	
	Progesterone	
	LH	
	FSH	
	ER-α SB	
	ER-β SB	
	AR SB	

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Supplemental Table VIII-25: Joint association of high ER-β serum bioactivity and gonadotrophins (top quantiles) with risk of breast cancer – cases that gave a sample less than 2 years before diagnosis.

SB - Hormones		Joint association of SB and gonadotrophins	
		ER-β SB	
		LH	FSH
		OR* (95% CI)	**OR* (95% CI)**
Not adjusted		0.476 (0.096-1.912) p=0.328	1.408 (0.261-3.492) p=0.566
		OR (95% CI)**	**OR** (95% CI)**
Adjusted	Oestradiol	0.467 (0.073-1.682) p=0.316	1.380 (0.375-4.147) p=0.589
	Oestrone	0.463 (0.072-1.678) p=0.312	1.402 (0.379-4.421) p=0.573
	Androstenedione	0.510 (0.079-1.849) p=0.377	1.521 (0.412-4.600) p=0.484
	Testosterone	0.456 (0.071-1.665) p=0.304	1.433 (0.385-4.370) p=0.551
	DHEAS	0.477 (0.075-1.718) p=0.330	1.419 (0.386-4.264) p=0.557
	SHBG	0.427 (0.067-1.545) p=0.264	1.219 (0.331-3.671) p=0.740
	Progesterone	0.469 (0.073-1.689) p=0.319	1.420 (0.386-4.273) p=0.557
	LH		1.503 (0.406-4.566) p=0.498
	FSH	0.478 (0.075-1.726) p=0.332	
	ER-α SB	0.473 (0.073-1.736) p=0.329	1.425 (0.384-4.331) p=0.556
	ER-β SB		
	AR SB	0.473 (0.074-1.704) p=0.324	1.402 (0.378-4.254) p=0.574

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Supplemental Table VIII-26: Joint association of high AR serum bioactivity and oestrogens (top quantiles) with risk of breast cancer – cases that gave a sample less than 2 years before diagnosis.

SB/Hormones		Joint association of SB and oestrogens AR SB Oestradiol
		OR* (95% CI)
Not adjusted		1.129 (0.379-2,997) p=0.816
		OR (95% CI)**
Adjusted	Oestradiol	
	Oestrone	1.216 (0.385-3.250) p=0.714
	Androstenedione	1.226 (0.391-3.245) p=0.700
	Testosterone	1.005 (0.319-2.659) p=0.992
	DHEAS	1.126 (0.363-2.934) p=0.820
	SHBG	1.092 (0.350-2.869) p=0.867
	Progesterone	1.095 (0.352-2.858) p=0.863
	LH	1.114 (0.358-2.904) p=0.836
	FSH	1.151 (0.370-3.006) p=0.788
	ER-α SB	1.236 (0.386-3.366) p=0.695
	ER-β SB	1.189 ().381-3.121) p=0.741
	AR SB	

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Supplemental Table VIII-27: Joint association of high AR serum bioactivity and androgens (top quantiles) with risk of breast cancer – cases that gave a sample less than 2 years before diagnosis.

SB - Hormones		Joint association of AR SB and androgens		
		AR SB		
		Androstenedione	Testosterone	DHEAS
		OR* (95% CI)	**OR* (95% CI)**	**OR* (95% CI)**
Not adjusted		1.881 (1.685-5.052) p=0.214	2.322 (1.056-6.640) p=0.067	0.724 p=0.614
		OR (95% CI)**	**OR** (95% CI)**	**OR** (95% CI)**
Adjusted	Oestradiol	1.880 (0.644-4.912) p=0.215	2.387 (0.922-5.810) p=0.060	0.776 (0.176-2.435) p=0.699
	Oestrone	1.710 (0.582-4.498) p=0.295	2.080 (0.800-5.067) p=0.115	0.613 (0.140-1.900) p=0.447
	Androstenedione		2.008 (0.760-4.987) p=0.141	0.679 (0.152-2.166) p=0.554
	Testosterone	1.367 (0.453-3.704). p=0.553		0.569 (0.127-1.836) p=0.393
	DHEAS	1.844 (0.626-4.877) p=0.234	2.300 (0.882-5.643) p=0.075	
	SHBG	1.690 (0.577-4.436) p=0.304	2.276 (0.876-5.565) p=0.077	0.648 (0.148-2.013) p=0.500
	Progesterone	1.784 (0.603-4.742) p=0.263	2.233 (0.856-5.472) p=0.085	0.660 (0.148-2.092) p=0.524
	LH	1.821 (0.624-4.760) p=0.239	2.288 (0.888-5.524) p=0.072	0.710 (0.162-2.198) p=0.594
	FSH	1.873 (0.639-4.922) p=0.220	2.323 (0.899-5.631) p=0.068	0.716 (0.163-2.221) p=0.603
	ER-α SB	2.039 (0.673-5.607) p=0.180	2.552 (0.949-6.523) p=0.054	0.783 (0.175-2.516) p=0.710
	ER-β SB	2.472 (0.812-6.915) p=0.092	2.865 (0.997-7.269) p=0.028	0.847 (0.190-2.707) p=0.799
	AR SB			

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Supplemental Table VIII-28: Joint association of high AR serum bioactivity and progesterone (top quantiles) with risk of breast cancer – cases that gave a sample less than 2 years before diagnosis.

SB - Hormones		Joint association of AR SB and progesterone AR SB Progesterone
		OR* (95% CI)
Not adjusted		1.200 (0.380-3.828) p=0.756
		OR (95% CI)**
Adjusted	Oestradiol	1.238 (0.339-3.3.656) p=0.718
	Oestrone	1.093 (0.298-3.227) p=0.880
	Androstenedione	1.005 (0.271-3.016) p=0.994
	Testosterone	0.869 (0.233-2.626) p=0.816
	DHEAS	1.158 (0.311-3.509) p=0.807
	SHBG	1.201 (0.328-3.558) p=0.757
	Progesterone	
	LH	1.180 (0.325-3.454) p=0.778
	FSH	1.187 (0.325-3.494) p=0.772
	ER-α SB	1.371 (0.361-4.300) p=0.609
	ER-β SB	1.399 (0.379-4.217) p=0.575
	AR SB	

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Supplemental Table VIII-29: Joint association of high AR serum bioactivity and sex hormone-binding globulin (top quantiles) with risk of breast cancer – cases that gave a sample less than 2 years before diagnosis.

SB/Hormones		Joint association of SB and SHBG AR SB SHBG
		OR* (95% CI)
Not adjusted		0.377 (0.048-2.019) p=0.355
		OR (95% CI)**
Adjusted	Oestradiol	0.424 (0.023-2.301) p=0.418
	Oestrone	0.414 (0.022-2.258) p=0.406
	Androstenedione	0.511 (0.027-2.873) p=0.531
	Testosterone	0.383 (0.020-2.115) p=0.369
	DHEAS	0.380 (0.021-2.022) p=0.359
	SHBG	
	Progesterone	0.367 (0.020-1.954) p=0.342
	LH	0.385 (0.0201-2.054) p=0.366
	FSH	0.378 (0.023-2.013) p=0.357
	ER-α SB	0.421 (0.021-2.294) p=0.416
	ER-β SB	0.386 (0.019-2.101) p=0.370
	AR SB	

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Supplemental Table VIII-30: Joint association of high AR serum bioctivity and gonadotrophins (top quantiles) with risk of breast cancer – cases that gave a sample less than 2 years before diagnosis.

SB - Hormones		Joint association of AR SB and gonadotrophins	
		AR SB	
		LH	FSH
		OR* (95% CI)	**OR* (95% CI)**
Not adjusted		0.773 (0.156-3.425) p=0.742	1.492 (0.491-4.145) p=0.459
		OR (95% CI)**	**OR** (95% CI)**
Adjusted	Oestradiol	0.770 (0.117-2.990) p=0.739	1.462 (0.459-3.999) p=0.482
	Oestrone	0.781 (0.118-3.055) p=0.754	1.522 (0.475-4.189) p=0.440
	Androstenedione	0.829 (0.125-3.238) p=0.811	1.762 (0.547-4.883) p=0.300
	Testosterone	0.762 (0.115-3.007) p=0.731	1.728 (0.536-4.796) p=0.317
	DHEAS	0.771 (0.117-2.989) p=0.740	1.510 (0.473-4.142) p=0.447
	SHBG	0.816 (0.123-3.233) p=0.798	1.493 (0.464-4.143) p=0.464
	Progesterone	0.759 (0.115-2.944) p=0.725	1.522 (0.475-4.189) p=0.440
	LH		1.599 (0.497-4.435) p=0.390
	FSH	0.786 (0.188-3.108) p=0.761	
	ER-α SB	0.706 (0.107-2.731) p=0.657	1.488 (0.466-4.079) p=0.463
	ER-β SB	0.845 (0.128-3.319) p=0.830	1.627 (0.506-4.517) p=0.373
	AR SB		

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Supplemental Table VIII-29: Joint association of high ER-α serum bioactivity and oestrogens (top quantiles) with risk of breast cancer – cases that gave a sample more than 2 years before diagnosis.

SB - Hormones		Joint association of ER-α SB and oestrogens	
		ER-α SB	
		Oestradiol	Oestrone
		OR* (95% CI)	**OR* (95% CI)**
Not adjusted		1.954 (0.930-5.545) p=0.133	1.868 (0.819-4.739) p=0.158
		OR (95% CI)**	**OR** (95% CI)**
Adjusted	Oestradiol		1.945 (0.770-4.543) p=0.136
	Oestrone	2.173 (0.850-5.175) p=0.088	
	Androstenedione	2.071 (0.803-4.976) p=0.113	2.035 (0.801-4.789) p=0.114
	Testosterone	1.670 (0.644-4.017) p=0.266	2.110 (0.826-4.995) p=0.099
	DHEAS	1.953 (0.775-4.550) p=0.133	1.868 (0.744-4.323) p=0.158
	SHBG	1.974 (0.780-4.623) p=0.129	1.850 (0.734-4.299) p=0.167
	Progesterone	2.054 (0.809-4.836) p=0.110	1.987 (0.785-4.651) p=0.125
	LH	1.887 (0.746-4.411) p=0.156	1.910 (0.757-4.448) p=0.147
	FSH	1.951 (0.772-4.559) p=0.135	1.866 (0.742-4.318) p=0.159
	ER-α SB		
	ER-β SB	1.783 (0.699-4.204) p=0.201	1.580 (0.604-3.825) p=0.326
	AR SB	1.881 (0.725-4.526) p=0.171	1.762 (0.687-4.170) p=0.212

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Supplemental Table VIII-30: Joint association of high ER-α serum bioactivity and androgens (top quantiles) with risk of breast cancer – cases that gave a sample more than 2 years before diagnosis.

SB/Hormones		Joint association of SB and androgens	
		ER-α SB	
		Androstenedione	DHEAS
		OR* (95% CI)	**OR* (95% CI)**
Not adjusted		2.146 (1.194-6.765) p=0.074	1.295 (0.482-3.864) p=0.627
		OR (95% CI)**	**OR** (95% CI)**
Adjusted	Oestradiol	2.335 (0.954-5.401) p=0.052	1.238 (0.392-3.303) p=0.688
	Oestrone	2.120 (0.878-4.798) p=0.079	1.273 (0.404-3.391) p=0.650
	Androstenedione		1.084 (0.338-2.939) p=0.882
	Testosterone	1.571 (0.618-3.767) p=0.322	0.946 (0.290-2.629) p=0.920
	DHEAS	2.173 (0.894-4.259) p=0.072	
	SHBG	2.076 (0.858-4.712) p=0.089	1.271 (0.402-3.402) p=0.653
	Progesterone	2.277 (0.927-5.279) p=0.061	1.261 (0.389-3.492) p=0.673
	LH	2.128 (0.879-4.838) p=0.079	1.317 (0.416-3.530) p=0.607
	FSH	2.125 (0.878-4.826) p=0.080	1.271 (0.401-3.407) p=0.654
	ER-α SB		
	ER-β SB	1.855 (0.740-4.375) p=0.168	1.120 (0.349-3.049) p=0.835
	AR SB	2.035 (0.824-4.726) p=0.107	1.159 (0.359-3.183) p=0.787

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Supplemental Table VIII-31: Joint association of high ER-α serum bioactivity and progesterone (top quantiles) with risk of breast cancer – cases that gave a sample more than 2 years before diagnosis.

SB - Hormones		Joint association of ER-α SB and progesterone
		ER-α SB
		Progesterone
		OR* (95% CI)
		1.667 (0.627-5.389) p=0.350
		OR (95% CI)**
Adjusted	Oestradiol	1.688 (0.519-4.762) p=0.344
	Oestrone	1.628 (0.506-4.517) p=0.373
	Androstenedione	1.253 (0.378-3.591) p=0.689
	Testosterone	1.082 (0.324-3.129) p=0.890
	DHEAS	1.750 (0.521-5.153) p=0.329
	SHBG	1.722 (0.533-4.800) p=0.322
	Progesterone	
	LH	1.655 (0.512-4.615) p=0.359
	FSH	1.635 (0.505-4.571) p=0.372
	ER-α SB	
	ER-β SB	1.424 (0.433-4.049) p=0.527
	AR SB	1.481 (0.443-4.306) p=0.490

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Supplemental Table VIII-32: Joint association of high ER-α serum bioactivity and progesterone (top quantiles) with risk of breast cancer – cases that gave a sample more than 2 years before diagnosis.

SB/Hormones		Joint association of SB and SHBG ER-α SB SHBG
		OR* (95% CI)
Not adjusted		2.650 (1.012-8.210) p=0.095
		OR (95% CI)**
Adjusted	Oestradiol	3.092 (0.989-9.184) p=0.043
	Oestrone	2.577 (0.760-7.954-) p=0.106
	Androstenedione	2.622 (0.766-8.185-) p=0.102
	Testosterone	2.379 (0.689-7.483) p=0.145
	DHEAS	2.651 (0.783-8.165) p=0.095
	SHBG	
	Progesterone	3.180 (1.018-9.440) p=0.038
	LH	2.863 (0.836-8.986) p=0.075
	FSH	2.659 (0.785-8.199) p=0.094
	ER-α SB	
	ER-β SB	2.480 (0.727-7.699) p=0.122
	AR SB	2.468 (0.714-7.777) p=0.129

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Supplemental Table VIII-33: Joint association of high ER-α serum bioactivity and gonadotrophins (top quantiles) with risk of breast cancer – cases that gave a sample more than 2 years before diagnosis.

SB - Hormones		Joint association of ER-α SB and gonadotrophins	
		ER-α SB	
		LH	FSH
		OR* (95% CI)	**OR* (95% CI)**
Not adjusted		0.459 (0.108-2.134) p=0.305	1.386 (0.470-4.884) p=0.580
		OR (95% CI)**	**OR** (95% CI)**
Adjusted	Oestradiol	0.448 (0.070-1.614) p=0.291	1.392 (0.381-4.123) p=0.576
	Oestrone	0.445 (0.069-1.604) p=0.287	1.354 (0.371-4.003) p=0.608
	Androstenedione	0.447 (0.070-1.619) p=0.291	1.464 (0.399-4.362) p=0.521
	Testosterone	0.472 (0.073-1.720) p=0.326	1.642 (0.445-4.929) p=0.407
	DHEAS	0.458 (0.072-1.649) p=0.304	1.389 (0.379-4.123) p=0.580
	SHBG	0.462 (0.072-1.667) p=0.310	1.388 (0.379-4.135) p=0.581
	Progesterone	0.456 (0.071-1.642) p=0.301	1.434 (0.390-4.283) p=0.544
	LH		1.768 (0.472-5.492) p=0.350
	FSH	0.467 (0.073-1.696) p=0.318	
	ER-α SB		
	ER-β SB	0.404 (0.063-1.471) p=0.236	1.254 (0.340-3.753) p=0.705
	AR SB	0.441 (0.069-1.590) p=0.282	1.351 (0.369-4.044) p=0.612

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Supplemental Table VIII-34: Joint association of high ER-β serum bioactivity and oestrogens (top quantiles) with risk of breast cancer - cases that gave a sample more than 2 years before diagnosis.

SB - Hormones		Joint association of ER-β SB and oestrogens	
		ER-β SB	
		Oestradiol	Oestrone
		OR* (95% CI)	**OR* (95% CI)**
Not adjusted		1.351 (0.630-4.455) p=0.536	1.482 (0.687-3.834) p=0.360
		OR (95% CI)**	**OR** (95% CI)**
Adjusted	Oestradiol		1.551 (0.626-3.502) p=0.311
	Oestrone	1.315 (0.465-3.243) p=0.573	
	Androstenedione	1.291 (0.450-3.248) p=0.606	1.580 (0.635-3.588) p=0.294
	Testosterone	1.109 (0.385-2.793) p=0.835	1.623 (0.650-3.704) p=0.269
	DHEAS	1.352 (0.479-3.335) p=0.535	1.483 (0.602-3.325) p=0.360
	SHBG	1.282 (0.452-3.172) p=0.611	1.386 (0.560-3.119) p=0.451
	Progesterone	1.424 (0.501-3.546) p=0.471	1.537 (0.621-3.470) p=0.321
	LH	1.338 (0.472-3.313) p=0.550	1.417 (0.573-3.185) p=0.420
	FSH	1.346 (0.476-3.328) p=0.542	1.480 (0.599-3.333) p=0.364
	ER-α SB	1.157 (0.402-2.904) p=0.769	1.181 (0.466-2.732) p=0.709
	ER-β SB		
	AR SB	1.366 (0.475-3.448) p=0.530	1.400 (0.559-3.199) p=0.444

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Supplemental Table VIII-35: Joint association of high ER-β serum bioactivity and androgens (top quantiles) with risk of breast cancer - cases that gave a sample more than 2 years before diagnosis.

SB/Hormones		Joint association of SB and androgens	
		ER-β SB	
		Androstenedione	DHEAS
		OR* (95% CI)	**OR* (95% CI)**
Not adjusted		2.133 (0.905-5.328) p=0.076	1.786 (0.633-4.588) p=0.251
		OR (95% CI)**	**OR** (95% CI)**
Adjusted	Oestradiol	2.164 (0.891-4.942) p=0.074	1.997 (0.670-5.361) p=0.185
	Oestrone	1.944 (0.770-4.540) p=0.136	1.898 (0.647-4.978) p=0.210
	Androstenedione		1.481 (0.491-4.000) p=0.456
	Testosterone	1.307 (0.494-3.211) p=0.570	1.346 (0.444-3.654) p=0.574
	DHEAS	2.189 (0.895-5.039) p=0.072	
	SHBG	1.973 (0.816-4.477) p=0.113	1.709 (0.586-4.429) p=0.291
	Progesterone	1.995 (0.779-4.741) p=0.129	2.088 (0.675-5.901) p=0.176
	LH	2.079 (0.860-4.714) p=0.088	1.791 (0.613-4.655) p=0.251
	FSH	2.122 (0.880-4.797) p=0.079	1.770 (0.609-4.576) p=0.259
	ER-α SB	1.693 (0.677-3.973) p=0.239	1.573 (0.524-4.227) p=0.387
	ER-β SB		
	AR SB	2.159 (0.872-5.041) p=0.082	1.608 (0.537-4.302) p=0.363

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Supplemental Table VIII-36: Joint association of high ER-β serum bioactivity and progesterone (top quantiles) with risk of breast cancer - cases that gave a sample more than 2 years before diagnosis.

SB - Hormones		Joint association of ER-β SB and progesterone
		ER-β SB
		Progesterone
		OR* (95% CI)
Not adjusted		1.063 (0.353-3.405) p=0.916
		OR (95% CI)**
Adjusted	Oestradiol	1.143 (0.313-3.344) p=0.821
	Oestrone	1.397 (0.438-3.801) p=0.535
	Androstenedione	0.922 (0.247-2.779) p=0.893
	Testosterone	0.740 (0.198-2.225) p=0.617
	DHEAS	1.083 (0.289-3.289) p=0.896
	SHBG	1.048 (0.290-3.009) p=0.936
	Progesterone	
	LH	1.034 (0.286-2.974) p=0.955
	FSH	1.045 (0.290-2.995) p=0.940
	ER-α SB	1.129 (0.351-3.090) p=0.823
	ER-β SB	
	AR SB	1.318 (0.406-3.663) p=0.616

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Supplemental Table VIII-37: Joint association of high ER-β serum bioactivity and gonadotrophins (top quantiles) with risk of breast cancer - cases that gave a sample more than 2 years before diagnosis.

SB - Hormones		Joint association of ER-β SB and gonadotrophins	
		ER-β SB	
		LH	FSH
		OR* (95% CI)	**OR* (95% CI)**
Not adjusted		0.769 (0.211-2.656) p=0.681	1.894 (0.637-5.702) p=0.248
		OR (95% CI)**	**OR** (95% CI)**
Adjusted	Oestradiol	0.751 (0.172-2.325) p=0.656	1.915 (0.589-5.430) p=0.241
	Oestrone	0.759 (0.174-2.349) p=0.668	1.880 (0.579-5.328) p=0.254
	Androstenedione	0.770 (0.175-2.397) p=0.685	1.850 (0.567-5.278) p=0.269
	Testosterone	0.718 (0.163-2.250) p=0.609	1.796 (0.547-5.151) p=0.296
	DHEAS	0.769 (0.176-2.378) p=0.682	1.893 (0.584-5.356) p=0.248
	SHBG	0.712 (0.163-2.211) p=0.598	1.709 (0.525-4.855) p=0.334
	Progesterone	0.756 (0.173-2.339) p=0.663	1.865 (0.574-5.281) p=0.260
	LH		2.153 (0.655-6.232) p=0.172
	FSH	0.779 (0.178-2.415) p=0.697	
	ER-α SB	0.648 (0.146-2.040) p=0.504	1.663 (0.505-4.855) p=0.365
	ER-β SB		
	AR SB	0.755 (0.172-2.340) p=0.662	1.822 (0.559-5.184) p=0.280

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Supplemental Table VIII-38: Joint association of high AR serum bioactivity and oestrogens (top quantiles) with risk of breast cancer - cases that gave a sample more than 2 years before diagnosis.

SB - Hormones		Joint association of AR SB and oestrogens	
		AR SB	
		Oestradiol	Oestrone
		OR* (95% CI)	**OR* (95% CI)**
Not adjusted		1.969 (0.825-4.877) p=0.129	2.388 (0.955-5.951) *p=0.058*
		OR (95% CI)**	**OR** (95% CI)**
Adjusted	Oestradiol		2.066 (0.762-5.132) p=0.130
	Oestrone	2.209 (0.864-5.264) p=0.081	
	Androstenedione	1.979 (0.774-4.696) p=0.133	2.520 (0.972-6.135) p=0.046
	Testosterone	1.775 (0.694-4.191) p=0.205	2.368 (0.858-6.042) *p=0.079*
	DHEAS	1.983 (0.785-4.626) p=0.125	2.411 (0.935-5.831) *p=0.056*
	SHBG	1.862 (0.736-4.255) p=0.165	2.336 (0.903-5.672) *p=0.066*
	Progesterone	1.944 (0.769-4.537) p=0.137	2.547 (0.980-6.245) p=0.045
	LH	1.964 (0.776-4.597) p=0.132	2.233 (0.865-5.408) *p=0.082*
	FSH	1.968 (0.778-4.598) p=0.130	2.395 (0.927-5.808) *p=0.059*
	ER-α SB	1.673 (0.639-4.064) p=0.27	1.914 (0.722-4.761) p=0.172
	ER-β SB	1.792 (0.702-4.223) p=0.197	2.040 (0.765-5.121) p=0.137
	AR SB		

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Supplemental Table VIII-39: Joint association of high AR serum bioactivity and androgenss (top quantiles) with risk of breast cancer - cases that gave a sample more than 2 years before diagnosis.

SB/Hormones		Joint association of SB and androgens
		AR SB
		DHEAS
		OR* (95% CI)
Not adjusted		1.607 (0.572-4.062) p=0.338
		OR (95% CI)**
Adjusted	Oestradiol	1.661 (0.574-4.257) p=0.312
	Oestrone	1.474 (0.517-3.689) p=0.431
	Androstenedione	1.437 (0.489-3.747) p=0.478
	Testosterone	1.249 (0.419-3.306) p=0.668
	DHEAS	
	SHBG	1.520 (0.529-3.852) p=0.400
	Progesterone	1.551 (0.524-4.088) p=0.394
	LH	1.580 (0.550-4.008) p=0.358
	FSH	1.582 (0.550-4.012) p=0.357
	ER-α SB	1.382 (0.466-3.635) p=0.530
	ER-β SB	1.370 (0.467-3.551) p=0.537
	AR SB	

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Supplemental Table VIII-40: Joint association of high AR serum bioactivity and sex hormone-binding globulin (top quantiles) with risk of breast cancer - cases that gave a sample more than 2 years before diagnosis.

SB/Hormones		Joint association of SB and SHBG AR SB SHBG
		OR* (95% CI)
Not adjusted		2.650 (1.012-8.210) p=0.095
		OR (95% CI)**
Adjusted	Oestradiol	3.092 (0.989-9.184) p=0.043
	Oestrone	2.577 (0.760-7.954-) p=0.106
	Androstenedione	2.622 (0.766-8.185-) p=0.102
	Testosterone	2.379 (0.689-7.483) p=0.145
	DHEAS	2.651 (0.783-8.165) p=0.095
	SHBG	
	Progesterone	3.180 (1.018-9.440) p=0.038
	LH	2.863 (0.836-8.986) p=0.075
	FSH	2.659 (0.785-8.199) p=0.094
	ER-α SB	
	ER-β SB	2.480 (0.727-7.699) p=0.122
	AR SB	2.468 (0.714-7.777) p=0.129

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Supplemental Table VIII-41: Joint association of high AR serum bioactivity and progesterone (top quantiles) with risk of breast cancer - cases that gave a sample more than 2 years before diagnosis.

SB - Hormones		Joint association of AR SB and progesterone AR SB Progesterone
		OR* (95% CI)
Not adjusted		2.305 (0.847-5.809) p=0.085
		OR (95% CI)**
Adjusted	Oestradiol	1.870 (0.636-4.930) p=0.223
	Oestrone	2.251 (0.822-5.694) p=0.095
	Androstenedione	1.884 (0.672-4.881) p=0.204
	Testosterone	1.360 (0.448-3.717) p=0.563
	DHEAS	2.703 (0.932-7.405) p=0.056
	SHBG	2.283 (0.830-5.809) p=0.091
	Progesterone	
	LH	2.309 (0.840-5.874) p=0.087
	FSH	2.274 (0.825-5.797) p=0.093
	ER-α SB	1.983 (0.695-5.271) p=0.179
	ER-β SB	2.031 (0.729-5.231) p=0.153
	AR SB	

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydro- epiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin

Supplemental Table VIII-42: Joint association of high AR serum bioactivity and gonadotrophins (top quantiles) with risk of breast cancer - cases that gave a sample more than 2 years before diagnosis.

SB - Hormones		Joint association of AR SB and gonadotrophins	
		AR SB	
		LH	FSH
		OR* (95% CI)	**OR* (95% CI)**
Not adjusted		0.412 (0.048-3.107) p=0.401	0.300 (0.043-2.631) p=0.249
		OR (95% CI)**	**OR** (95% CI)**
Adjusted	Oestradiol	0.385 (0.021-2.069) p=0.368	0.305 (0.017-1.562) p=0.255
	Oestrone	0.411 (0.022-2.200) p=0.400	0.302 (0.016-1.550) p=0.252
	Androstenedione	0.394 (0.021-2.123) p=0.379	0.321 (0.018-1.657) p=0.278
	Testosterone	0.396 (0.021-2.150) p=0.383	0.339 (0.018-1.753) p=0.302
	DHEAS	0.416 (0.022-2.226) p=0.406	0.297 (0.016-1.522) p=0.245
	SHBG	0.433 (0.023-2.340) p=0.429	0.302 (0.016-1.561) p=0.253
	Progesterone	0.399 (0.022-2.139) p=0.385	0.301 (0.016-1.547) p=0.251
	LH		0.332 (0.018-1.718) p=0.292
	FSH	0.422 (0.023-2.270) p=0.415	
	ER-α SB	0.309 (0.017-1.661) p=0.268	0.268 (0.015-1.387) p=0.209
	ER-β SB	0.364 (0.020-1.972) p=0.341	0.266 (0.014-1.379 p=0.207
	AR SB		

*OR values for top-bottom classification were based on controls only and were age adjusted. **OR values for SB adjusted for other hormones or SB treated as categorical variable, using top bottom classification and age adjusted. OR with p-values ≤0.05 marked with bold.

AR=androgen receptor; CI=confidence interval; DHEAS=dehydroepiandrosterone sulphate; ER=oestrogen receptor; FSH=follicular stimulating hormone; LH=luteinising hormone; OR=odds ratio; SB=serum bioactivity; SHBG=sex hormone-binding globulin