MASTER OF SCIENCE

Assessing risk of familial breast cancer effectiveness of current UK guidelines

Littlejohn, Lucy

Award date:
2017

Awarding institution:
University of Dundee

Assessing risk of familial breast cancer: effectiveness of current UK guidelines

Lucy Littlejohn

MSc by Research in Medicine
College of Medicine, Dentistry and Nursing
University of Dundee
March 2017

Contents

Acknowledgements

I would firstly like to thank my supervisor Dr Jonathan Berg for always being encouraging and available at every stage of the study. The guidance he provided was invaluable and I am grateful for his useful comments and engaging discussion. I would like to acknowledge Dr David Goudie from the clinical genetics department for helping with data availability at various stages in this research. My sincerest gratitude also goes to Jim Gibbs and the Pathology Department at Ninewells hospital for being so willing to help and share information from their databases. I also wish to thank the recently retired Simon Ogston from the University of Dundee, Department of Population Health Sciences for advice on statistical analyses. Finally, my thanks to family and friends for their patience and support throughout the process of completing this work.

Declaration

I declare that the content of this report is my own work and has not previously been submitted for any other assessment. The report is written in my own words and conforms to the University of Dundee's Policy on plagiarism and academic dishonesty. Unless otherwise indicated, I have consulted all of the references cited in this report.

Lucy Littlejohn Date

Summary

Breast cancer risk is a common indication for referral to clinical genetics. National Institute of Health and Care Excellence (NICE) guidelines use family history to stratify patients by 10-year risk of breast cancer from the ages 40-49. Patients are divided into low (10-year risk <3%), moderate (3-8%) and high risk (>8%). Those with a *BRCA* mutation are considered to be very high risk. Women at moderate or high risk are offered screening from age 40. This study aimed to assess the effectiveness of NICE risk categorisation at identifying women at risk of early onset breast cancer.

Family history data was obtained for unaffected women with a family history of breast cancer, aged <50 years, referred to Tayside clinical genetics from 2000-2010. Patients were risk stratified *de novo* by NICE criteria. Those who went on to develop breast cancer were identified.

1,409 women were included in the cohort, with a total of 15,414 patient-years of follow up. Of these patients, 35.84% were NICE low risk, 37.04% moderate risk and 27.11% were high risk. 22 *BRCA* mutation carriers were identified.

30 invasive breast cancers developed, 13 in moderate and 13 in high risk women. Kaplan-Meier analysis demonstrated no significant difference in cancer rates between low and moderate risk women from ages 40-49 (Log rank p=0.431). There was a significant difference from 40-49 years between low and high risk women (p=0.036), but not on exclusion of *BRCA* carriers (p=0.136). The 10-year absolute risk from 40-49 years was 0.82% (95% CI, 0.72-0.94%) for low risk, 1.68% (1.53-1.83%) for moderate risk, and 3.56% (3.34-3.80%) for high risk women. NICE absolute 10-year risk thresholds for screening were not met in any group.

This study provides some evidence that screening prior to age 50 in those without a *BRCA* mutation may be unnecessary. In the study cohort, NICE family history criteria identify women at increased risk of breast cancer, but not at the absolute risk thresholds suggested for screening. There is a need for further evaluation of NICE criteria.

List of Abbreviations

ANOVA, analysis of variance;

AUROC, area under the receiver operating curve;

BCAC, Breast Cancer Association Consortium;

BOADICEA, Breast and Ovarian Analysis of Disease Incidence and Carrier Estimation Algorithm;

CBC, contralateral breast cancer;

CI, confidence interval;

COCP, combined oral contraceptive pill;

DCIS, ductal carcinoma *in situ*;

ER, oestrogen receptor;

FDR, first degree relative;

FGFR2, fibroblast growth factor receptor 2;

FH, family history;

GWAS, genome-wide association study;

HER2, human epidermal growth factor 2;

HRR, homologous recombination repair;

HR, hormone receptor;

HRT, hormone replacement therapy;

ICER, incremental cost-effectiveness ratio;

IR, ionising radiation;

KM, Kaplan-Meier;

LCIS, lobular carcinoma *in situ*;

LD, linkage disequilibrium;

MAF, minor allele frequency;

MD, mammographic density;

MMSS, modified Manchester scoring system;

MMS, modified Manchester score;

MSS, Manchester scoring system;

NBSP, National Health Service breast screening programme;

NHEJ, non-homologous end-joining;

OR, odds ratio;

PR, progestogen receptor;

ROC, receiver operating curve;

RR, relative risk;

SD, standard deviation;

SDR, second degree relative;

TERT, telomerase reverse transcriptase;

TF, transcription factor;

TNBC, triple negative breast cancer.

List of Figures

List of Tables

1. Introduction

1.1. Overview of breast cancer

1.1.1. Breast cancer epidemiology and population screening

In the UK, breast cancer is one of the most common cancers in females, affecting up to 1 in 8 women in their lifetime (1). Risk is dependent on age, environmental and non-modifiable risk factors. Fortunately, treatment outcomes remain largely positive due to advances in the understanding of breast cancer, and population screening, with survival rates of >90% for women aged 40-69 (2). Over the last 20 years, rates of both invasive and *in situ* breast cancer have been increasing, however mortality rates continue to improve (see *Figure 1*). Detailed age dependent UK incidence and outcome data for breast cancer can be seen in *Table 1*. This demonstrates a peak incidence rate of invasive breast cancer in the older age group, however it is worth noting that the largest number of breast cancers occur in the middle age ranges. Therefore, the bulk of the UK breast cancer burden comes from women in their middle age, with peak numbers diagnosed between the ages of 60-69. In males, breast cancer still remains very uncommon with just 340 cases diagnosed in the UK in 2013 (1).

UK population screening programmes have been in place for breast cancer since 1988. Women between the ages of 50-70 are invited for mammographic screening every 3 years as part of the National Health Service Breast Screening Programme (NBSP), although studies are ongoing investigating the possibility of extending this to ages 47-73 (3). A comprehensive review by the Independent UK Panel on Breast Cancer Screening found that it has resulted in a 20% relative risk (RR) reduction in breast cancer mortality in the screened population (4). They also identified issues around over-diagnosis, however deemed that any harm this may cause to individual women through investigations and treatment, is largely outweighed by the 1,300 breast cancer deaths probably prevented every year (4). Despite the success of the NBSP, evidence for its cost effectiveness is less clear. A study which followed up 364,500 women for 35 years found a cost of £20,800 per quality adjusted life year gained associated with screening, slightly higher than the guidance threshold of

£20,000 set by NICE (5). The authors concluded only a moderate cost effectiveness of the screening programme. Despite this, the NBSP may be partly accountable for the both the increasing incidence and decreasing mortality seen in the UK population since the late 1980s.

Figure 1. Breast cancer trends over time for UK females. Data from cancerresearchuk.org (1)

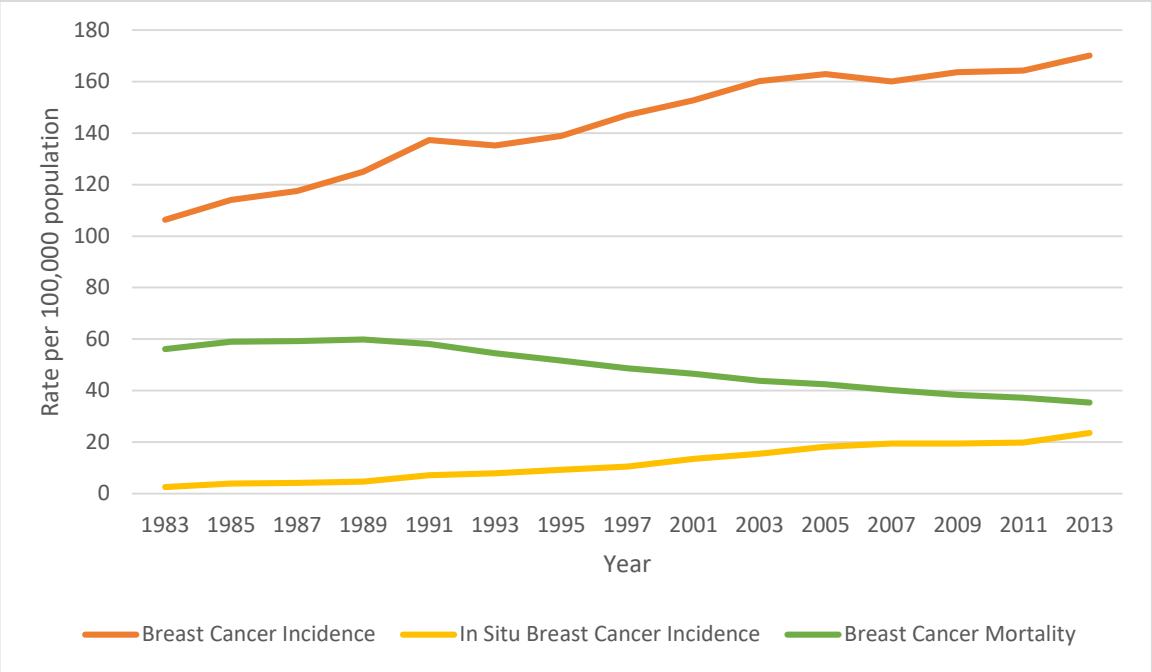


Table 1. UK incidence and mortality data by age range. Absolute number and incidence rates for 2013, mortality rates based on data from 2009-2014. Data obtained from cancerresearch.org (1)

Age Range (years)	Average number of cases per year	Rate per 100,000 (UK population)		
		Invasive breast cancer incidence	*In Situ* breast cancer incidence	Breast cancer mortality
20-24	34	1.6	0.1	0.1
25-29	206	9.5	0.7	0.9
30-34	586	27.5	1.8	3.1
35-39	1,277	62.7	4.9	7.2
40-44	2,804	121.4	11.7	14.9
45-49	5,189	219.1	32.5	24.4
50-54	5,934	277.6	61.9	35.1
55-59	5,101	273.3	47.2	44.6
60-64	6,541	351.2	56.1	55.5
65-69	6,951	409.3	62.7	66.5
70-74	4,513	343.8	37.3	84.5
75-79	4,405	393.0	24.5	115.2
80-84	3,800	425.1	17.4	160.0
85-89	2,683	453.8	14.6	222.8
>90	1,662	450.5	10.2	345.2
Overall	51,691	159.6	20.2	35.2

1.1.2. Pathology

Breast cancer is categorised broadly according to site of origin. This splits the disease into ductal carcinoma, occurring in the mammary ducts, and lobular carcinoma, occurring in the breast lobular tissue i.e. the milk producing tissue of the breast. In addition, this is further sub-categorised into ductal/lobular carcinoma *in situ* (DCIS/LCIS) and invasive cancers, dependent on whether or not the tumour has invaded into the surrounding structures. Lobular carcinoma is thought to account for approximately 10% of invasive breast cancers, and be more strongly associated with exposure to female hormones (discussed later), than ductal carcinoma (6, 7). Although both cancers seem to be somewhat linked with a familial predisposition, lobular carcinoma has been found to be less prevalent in groups of people with certain high risk mutations i.e. *BRCA1* and *TP53* (discussed later), which cause a genetic predisposition to breast cancer (7). Although this broad classification is still used clinically, many other ways of classifying breast cancer pathologically have been developed, and have found to be relevant to both the treatment and aetiology of breast cancer.

Immunohistochemistry allows for the identification of subtypes of breast cancer based on receptor expression. This includes oestrogen receptors (ER), progestogen receptors (PR) and human epidermal growth factor 2 receptors (HER2). The ER and PR are often grouped together as 'hormone receptors' (HR), as PR expression is considered to be an indicator of ER signalling (8). However, it has been suggested that absence of PR expression may be a useful prognostic indicator, shown to influence recurrence free and overall survival (9, 10). ER positivity has been found to be present in approximately 75% of breast cancers, and is more common in post-menopausal breast cancer (11). However, in individuals with highly deleterious mutations in *BRCA1*, a breast cancer predisposition gene, around 75% of tumours are found to be ER negative (12). Like the general population, 75% of tumours in those with a mutation in *BRCA2* (discussed later) are found to be ER positive (12). ER receptor status also has implications for treatment, as these tumours are sensitive to treatment with anti-oestrogens, such as tamoxifen which is (primarily) an oestrogen

antagonist, and aromatase inhibitors which prevent the synthesis of oestrogen peripherally (11). Approximately 20% of breast tumours are HER2 positive, and expression of this receptor was originally linked to more aggressive tumours (13). However, since the advent of therapies which target the HER2 receptor specifically, the prognosis for this subtype has improved. Trastuzumab (Herceptin) is a HER2 specific receptor modulator which prevents it's downstream signalling to oncogenic proteins (8). HER2 positivity has been found to be very rare in *BRCA* mutation carriers, with just 2.1% of women with *BRCA1* and 6.8% of *BRCA2* related breast tumours being positive (14). Another subgroup are 'triple negative' breast cancers (TNBC) which express neither HR nor HER2. They are seen more often in younger women and are associated with a more aggressive phenotype, with no targeted therapies like those available for HR or HER2 positive breast cancer (15). 65-80% of breast cancers in *BRCA1* mutation carriers have been found to be TNBC and both *BRCA1* and *BRCA2* mutation carriers have been found to present with more advanced TNBC than non-carriers (15).

The above classifications are used clinically as they have the ability to inform treatment. However, another classification system known as the 'intrinsic subtype', which is not currently being used clinically, is being extensively investigated. These subtypes which were identified through unique gene-expression profiles are intrinsically linked to receptor status, but have been found to better differentiate survival outcomes, and may potentially lead to differing therapeutic strategies (16). The names and some of the features of the five intrinsic subtypes are outlined in *Table 2.* As well as different gene-expression and prognostic outcome, it has been possible to differentiate levels of genomic instability, ploidy, methylation profiles and protein expression between the groups. For example, the DNA of basal-like tumours tends to be hypo-methylated, whereas luminal B tumours are hyper-methylated; Luminal A tumours are mainly diploid with little genomic instability, compared to mainly aneuploid, highly genomically unstable HER2 and basal-like tumours (17). In general the tumours have also been mapped to an immunohistochemical profile. The exceptions are the normal-like tumour (which is similar to luminal A), and luminal B which seems to present as either HER2 positive or negative (16). Family history (FH)

of breast cancer in a first degree relative (FDR) has found to be linked to luminal A, luminal B and HER2-overexpressing breast cancers (18). In addition, reduced expression of *BRCA1* was found to be linked to basal-like breast cancer. This is perhaps unsurprising, as it has a TNBC expression profile, seen more commonly in individuals with *BRCA1* mutations (19). Research in the field of gene-expression classification is extensive and it is clear that it may greatly impact clinical care in the future. However, as yet current practise uses an immunohistochemical-based classification.

Table 2 . Intrinsic breast cancer subtypes - adapted from Dai X et al. (15).

Intrinsic Subtype	Immunohistochemistry profile			Outcome	Prevalence
	ER	PR	HER		
Luminal A	+	+	-	Good	23.7%
Luminal B	+	+	+	Intermediate	38.8%
			-	Poor	14%
HER2 over-expression	-	-	+	Poor	11.2%
	-	-	-	Poor	12.3%
Basal	+	+	-	Intermediate	7.8%
Normal-like					

1.2. Breast Cancer Risk Factors

1.2.1. Socioeconomic status and ethnicity

As opposed to many cancers, breast cancer is widely considered to be a disease primarily affecting socioeconomically affluent populations, with consistently reported increased incidence in higher social groups (20)**.** However, despite this, a study of a Scottish population from 1994 demonstrated that mortality was greater in lower social groups, with 66% 5-year survival in the most affluent group and 55% 5-year survival in the least affluent groups, despite no significant differences in tumour biomarkers (21). This is a trend which has been replicated in more recent studies, and in those of different study populations. Additionally, the evidence is growing that women of lower socioeconomic status are more likely to have breast cancers with a poorer prognosis and which present at a later stage (20, 22, 23). Literature reviews have noted the strongest associations between social status and breast cancer incidence appear to apply to post-menopausal women. This is thought to reflect the fact that post-menopausal breast cancers, in general, may be more likely to be associated with the cumulative effect of specific lifestyle factors, whilst younger breast cancers may be more tied to genetics and less reliant on factors such as social class (20)**.** Most studies investigating socioeconomic status and breast cancer have been based on the assumption that social class itself is unlikely to have a direct effect on breast carcinogenesis, and that it is instead a marker for a sum of behaviours and exposures which encompass a profile of risk. However, it is becoming more evident that the discrepancy is difficult to explain by this model, and thus, the effect is established but as of yet poorly understood (20).

Much of the work concerning race and breast cancer has been carried out in populations from the USA, however the results have been more or less consistent. It is well recognised that breast cancer as a whole tends to affect white women more so than other racial groups (24, 25). However, there appears to be crossover, with rates higher in black women than white women at age <40 years – one study demonstrated incidence rates of 16.8/100,000 and 15.1/100,000 for black and white women respectively, aged <40 years (26)**.** Racial group also has a significant impact on breast cancer survival outcomes. Figures from the USA have shown a 5-year

survival of 91.4% for Asian women, 88.6% for non-Hispanic white women, 87% for Hispanic/Latina women, 85.4% for Native American women and 78.9% for African American women (27)**.** African American women have also been consistently shown to present with more aggressive breast cancer of ER-ve/PR-ve and TNBC subtypes, which have fewer treatment options and a poorer prognosis (24, 26). It has been suggested that the poorer survival outcomes in ethnic minority groups may be linked to poorer access to healthcare and less screening. African American women have for some time been screened at least as often as white women in the USA, with no change in prognosis, however it is suggested that there may be significant over-reporting of mammography attendance in this group (24, 28). There is increasing evidence that tumour behaviour and biology may vary by ethnicity, meaning that the observed differences in prognosis may be explained by biological mechanisms – further research into these differences is needed to improve understanding and treatment of breast cancers in different ethnic groups (24, 29).

1.2.2. Hormonal and reproductive factors

Steroid hormones, more specifically sex hormones such as oestrogen are known to have significant effects on growth, differentiation and functioning of the breast, and are major promoters of proliferation in both normal and neoplastic breast tissue (30, 31). It is therefore somewhat unsurprising that both exogenous and endogenous hormonal factors have been shown to have an association with breast cancer pathogenesis. These factors are largely linked to oestrogen exposure, which is considered to be the main modifiable risk factor for breast cancer (1).

Endogenous

Breast cancer has been found to be more common in women who have never breast fed, those with early age at menarche and later menopause, and nulliparous women (32)**.** An earlier age at menarche and later menopause is thought to contribute to breast cancer risk due to prolonged exposure of the breast epithelium to sex hormones, such as oestrogen and progestogen. This is due to an increased number of ovulatory cycles throughout life (33)**.** Accordingly, breast cancer risk has been shown to increase by 17% for every 5-year increase in the age at menopause (34)**.**

Similarly, the decrease in breast cancer risk associated with increased duration of breast feeding is thought to be linked to the reduced hormone exposure resulting from breast-feeding, which prolongs the anovulatory period following childbirth (33). What is less clear is the cause of the association between breast cancer and parity. Later age at first full-term pregnancy and lower/nulli-parity are both well-established factors which increase breast cancer risk. However, interestingly, it has been shown that uniparous women (particularly those aged 30 or more) have an elevated risk of breast cancer soon after delivery compared to nulliparous women - this declines some years later (35). Generally, circulating sex hormones are likely to be higher in pregnancy. Therefore, it is hypothesized that risk decrease may be due to the altered hormonal environment specific to pregnancy, maturation of breast tissue in response to pregnancy (resulting in a greater degree of differentiation), an effect on mammary stem cells, or changes in oestrogen responses of the mammary gland (35). The mechanism is poorly understood however it is clear that the protective effect of parity is complicated and dependent on several additional factors.

As with many other cancers, increased age is also an established breast cancer risk factor, and it is thought one reason for this is the effect of endogenous hormones. Although the levels of circulating oestrogens in post-menopausal women are greatly decreased compared to pre-menopausal women, the difference in the composition of breast tissue may contribute to increased risk with age. Breast stromal cells and specifically adipocytes in breast tissue contain aromatase enzymes which convert circulating cholesterol to oestradiol (32), a potent naturally occurring oestrogen. In post-menopausal women, the proportion of breast tissue made up of adipocytes as opposed to fibrous tissue is much greater than in pre-menopausal women. Due to this, it is recognised that oestradiol levels in the breast tissue of post-menopausal women are much greater than plasma levels (36).

Exogenous

Long term hormone replacement therapy (HRT) with oestrogen-containing preparations is also thought to contribute significantly to the increased risk of breast cancer in post-menopausal women. There is a reported RR of 1.43 for post-

menopausal women who have ever used HRT, and 1.66 among current users, increasing with total duration of use (37). UK guidelines suggest that HRT is inappropriate for women who have breast cancer, or who are at high risk due to a strong FH, and alternatives for managing menopausal symptoms should be sought (38). If symptoms are unmanageable and the patient is informed of the risk, HRT should be used at as low a dose for as little a time as possible, and these women should be preferentially given oestrogen-only preparations. The increased risk is of course applicable to any woman taking HRT, and it comes with other risks such as increased incidence of venous thrombo-embolism. However, HRT is an effective treatment for menopausal symptoms such as vasomotor flushing, mood disturbance and sexual dysfunction, as well as longer term benefits including improved bone mass density, and decreased risk of fragility fractures and coronary heart disease (38). Therefore, whilst the increased risk of breast cancer associated with HRT is not insignificant, as with any treatment intervention the risks and benefits to the individual patient should be assessed.

Perhaps more relevant to pre-menopausal women is the impact of oestrogen-containing contraceptive preparations such as the combined oral contraceptive pill (COCP). A review and analysis of forty-four studies evaluating COCP use and breast cancer incidence found that having ever used a COCP, versus having never used one, had a significant impact on risk of breast cancer (odds ratio (OR) =1.08 (95% confidence interval (CI), 1.00-1.17)). Risk was further increased for current users, or those with a shorter time since last use, however no relationship was found with total duration of use (39). As with HRT, the COCP is contraindicated in women with breast cancer, history of breast cancer or at high risk of breast cancer according to UK guidelines (40). An interesting caveat is that having ever used the COCP is associated with a 30% reduction in the risk of ovarian cancer in the general population, and has also been found to reduce ovarian cancer risk in women who carry *BRCA* mutations (41), which hugely increase lifetime risk of developing breast and ovarian cancer. Meta-analysis has shown an inverse relationship between ovarian cancer and COCP use in *BRCA* mutation carriers, and a modest but not statistically significant increase in breast cancer risk in this group (though absolute risk would remain much greater

than the general population) (41). It is therefore suggested that use of the COCP in this specific high risk population, may not be contraindicated due to the net benefit of reduced risk of ovarian cancer; this helps avoids definitive surgery such as oophorectomy in women who have not yet completed their family.

Overall, HRT and the COCP use is not recommended in women with breast cancer or with an increased risk. Any woman seeking these treatments should be made aware of the potential risks, benefits and counselled accordingly, though for most, the increased risk of breast cancer is not a contraindication.

Reproductive and hormonal risk factors and triple negative breast cancer

Although the hormonal risk factors discussed above are well established for breast cancer overall, the effect of these differs when looking at different breast cancer subtypes. Evidence for associations with reproductive risk factors has been found to be most consistent for tumours which are ER +ve and/or PR +ve. Associations are less convincing for TNBC, a subtype which accounts for 10-20% of diagnosis (42). Since oestrogen is a major promoter of proliferation in neoplastic breast tissue, it is unsurprising that ER +ve tumours may proliferate in response to oestrogen exposure – for TNBC however, the picture is less clear. It has been shown that reproductive risk factors are more strongly associated with HR +ve tumours than with ER-ve and PR-ve tumours (31, 43), although this is poorly understood. HER2 status appears to be less of a discriminator when it comes to risk factors (31). It has been suggested that TNBC may represent a distinct subtype of breast cancer epidemiologically. Differences that have been established include an association between increased parity and increased TNBC risk and increased age at first pregnancy and decreased TNBC risk, quite contrary to traditional perceptions of breast cancer risk factors (42). TNBC also has a tendency to affect women of a younger age, and is more common in black women compared to HR +ve cancer (31). The epidemiology of TNBC specifically is poorly understood, and further work is required to establish the relevance of traditional breast cancer risk factors in this subtype.

1.2.3. Modifiable Lifestyle Factors

There is well-established evidence that modifiable factors contribute to the pathogenesis of multiple cancers, including breast cancer, although the biological mechanisms are not always well described. Nonetheless, the factors discussed below have been associated with breast cancer, and therefore represent potential lifestyle changes that may confer a modest reduction in breast cancer risk.

Obesity, diet and physical activity

Obesity has been found to increase breast cancer risk in post-menopausal women. In a case-control study from the USA, there was an increased risk of 8% for every 5kg heavier a woman is compared to her lowest adult weight (44)**.** However, for younger women the relationship is less clear, with some suggesting that obesity pre-menopause may in fact decrease the risk of getting breast cancer pre-menopause, or at least have negligible effect (45, 46)**.** It is speculated that this may be as a result of the anovulatory state that can occur in obese pre-menopausal women, resulting in less endogenous oestrogen exposure.

Under normal physiological circumstances, circulating sex hormones in post-menopausal women would be expected to be lower than that of pre-menopausal women, which would be expected to have a risk-lowering effect. However, in obesity there is an increase in peripheral adipose tissue which express aromatase enzymes responsible for production of oestradiol. Additionally, it has been demonstrated that levels of *CYP19A1* mRNA (which encodes aromatase) levels are 2-4 times higher in fat from the buttocks, thighs and abdomen of post-menopausal women compared to young women (47)**.** Therefore, the increased risk associated with obesity may in part be due to higher circulating oestrogens in obese post-menopausal women (48)**.** Another causal link may be due to a metabolic state which often arises to some degree in obesity known as ‘metabolic syndrome’. This involves insulin resistance, dysglycaemia or hyperglycaemia, dyslipidaemia, hypertension and an overall chronic inflammatory response, which together is thought to produce a pro-carcinogenic state. This may link obesity to breast cancer, as well as other cancers (48). Accordingly, weight loss of 5% has been shown to reduce breast cancer risk by as

much as 25-40%, and is of probable benefit in all overweight women regardless of the presence of other risk factors (49).

Since, in basic terms, weight is dependent on net calories consumed vs calories used, it is intrinsically linked to both diet and exercise. Whilst weight loss generally has been linked to a decreased risk of breast cancer, evidence is less certain when it comes to the specifics of either diet, or physical activity in particular. Research has been carried out regarding dietary composition and its effect on breast cancer risk. Specific dietary components have been linked to breast cancer, for example, as with many cancers, increased fruit and vegetable intake decreases risk (33). There has been shown to be a 5% decrease in risk for every additional 10g of fibre consumed per day, particularly soluble fibres (50). High fat intake, particularly from unsaturated fat has been linked to breast cancer risk, and one study demonstrated that a reduced fat diet significantly improves relapse free survival in breast cancer patients (48, 51). This was not however replicated in further studies (52), so the evidence remains uncertain. Consumption of 'well-done' red meat, and processed meat has also been linked to a significantly increased risk of breast cancer, although it is unclear as to whether or not this is due to the fat content of meat, chemical exposure from processing or if this is just a marker of unhealthy lifestyle generally (33, 49). Much hype surrounded the possibility of dairy products as a contributor to breast cancer, as it was hypothesized that oestrogen consumption from this source would increase risk. However, this has been dispelled and in fact it has been shown that that there is a decreased rate of breast cancer in those who consume more dairy products (49). Products derived from soya have also caused controversy, as they are high in isoflavones which have been shown to have both weak oestrogenic and anti-oestrogenic effects as well as inhibiting cancer growth in laboratory conditions (49). Despite the weak oestrogenic effect, intake of 5g of soya protein/day is associated with a reduction in breast cancer risk in Asian populations, although it is thought that this impact may be as a result of effects on the developing breast in which case intake of soya products as a child/adolescent may be more important (49, 53).

Although physical activity will contribute to weight loss, it is less clear whether or not it itself reduces risk of breast cancer. Physical activity is associated with a reduction in endogenous sex hormones, reduction in insulin resistance and reduction in chronic inflammatory processes, which overall could contribute to reduced risk (49). It has been shown to reduce risk in both pre- and post-menopausal women, however the effect was limited in women who were overweight or obese (54). This finding possibly gives strength to the assertion that the main benefit of exercise in the reduction of breast cancer risk is by decreasing total body weight rather than a direct effect of exercise itself. Overall, maintaining a healthy weight, or losing weight through exercise or a calorie-reduced diet reduces the risk of breast cancer, particularly in post-menopausal women. However, the evidence so far is limited as to whether there are any diet-specific or exercise-specific risk-reducing measures which are beneficial.

Alcohol

Alcohol consumption has been consistently associated with an increased risk of developing breast cancer in both pre- and post-menopausal women, with an overall RR of 1.6 (33). For every 10g/day increase in alcohol consumption there has been found to be a 7-10% increase in risk of breast cancer, with 4-10% of breast cancer diagnosed in the USA attributable to alcohol consumption (55). There has been speculation as to the biological link between the two. Despite alcohol being largely metabolised by the liver, it is known that human breast tissue has a modest capacity to do so, and it may be that metabolites resulting from this have a direct carcinogenic effect. Acetaldehyde and reactive oxygen species, both products of alcohol metabolism are known to cause DNA point mutations, crosslinks, chromosomal aberrations and strand breaks (55). Accumulation of these abnormalities with continued exposure could contribute to neoplastic change within breast tissue. Additionally, another effect of alcohol is to increase circulating levels of oestrogen, which is thought to contribute to propagation of breast cancer rather than neoplastic change itself. In a study in which 30g/day of ethanol was consumed by women age 21-40 years for 3 menstrual cycles, it was shown that there was a 28% increase and 21% increase in plasma oestradiol and oestrone (endogenous oestrogens)

respectively (56) - an increase in oestrogens has also been demonstrated in post-menopausal women after alcohol consumption (57). It is thought that this may contribute to proliferation of already neoplastic breast tissue, especially given that ethanol has been shown, by some mechanism, to cause proliferation of ER+ve but not ER-ve breast cancer cell lines (58).

Smoking

Tobacco smoking has been indicated as cause or risk factor for the development of several human cancers including cancer of the upper airways, oral cavity, oesophagus, liver, pancreas, and bladder, with the highest RR for current smokers being that of lung cancer (RR=8.96) (59). The link with breast cancer however is less clear. The role of smoking in breast cancer seems to be of particular relevance for ER +ve breast cancers in people with specific polymorphisms in the *NAT2* gene, involved in the metabolism of tobacco products (60). Additionally, the risk appears to be greater for those who began smoking in adolescence, with an OR of 1.5 (95% CI, 0.9-2.5) for women who started smoking between ages 10-14 years compared to 1.2 (0.8-1.5) for those who began aged >20 years (61). Overall, the evidence would suggest that tobacco smoking is associated with a modest increase in risk of pre-menopausal breast cancer with some reports of an association with post-menopausal breast cancer as well (48, 49).

1.2.4. Mammographic Density

Mammographic Density (MD) refers to the component of tissue on breast mammography which appears radio-dense. Radio-dense tissue represents fibroglandular components of the breast, and radio-transparent areas represent fatty tissue. An increased MD has been consistently linked to an increased risk of breast cancer in both pre- and post-menopausal women. Studies investigating the relationship between MD and breast cancer risk have taken several approaches, including measuring absolute dense area, percentage dense area and absolute non-dense area as measurements of MD. In a meta-analysis (62), the age-adjusted OR in pre-menopausal women, for one standard deviation increment in absolute dense area was 1.38 (95% CI, 1.3-1.49), for percentage dense area OR=1.45 (1.35-1.55) and for absolute non-dense area 0.78 (0.69-0.89). For post-menopausal women, the same figures were 1.37 (1.33-1.40), 1.53 (1.44-1.64) and 0.79 (0.73-0.85) respectively (62)**.** For post-menopausal women, percentage dense area and absolute non-dense area were corrected for BMI and parity, however the same correction did not significantly affect absolute dense area or the figures quoted for pre-menopausal women. The authors of this meta-analysis concluded that absolute and percentage dense area were both strong risk factors for breast cancer, with percentage dense area being the stronger of the two risk factors. Evidence for the value of absolute non-dense area as a risk determinant was deemed inconclusive. Greater associations have been drawn when comparing the extremes of MD, with one meta-analysis quoting a RR of breast cancer of 4.64 (95% CI 3.64-5.91) for dense area >75%, compared to <5% (63)**.** The confounding effect of BMI in post-menopausal women is an interesting component of analysis of MD and breast cancer risk. As discussed previously, higher BMI is known to increase breast cancer risk in post-menopausal women, partly due to increased adipose tissue both peripherally and in the breast, contributing to increased aromatisation of sex hormone precursors. However, increased MD, supposedly representing decreased breast adiposity, is also linked to an increased breast cancer risk. An explanation which has been suggested for this discrepancy is that it is largely primary stromal preadipocytes which have aromatase activity. This activity may diminish as they differentiate into mature adipocytes (64), which may be present in fatty radiotranslucent breast tissue.

The biological mechanisms behind increased breast cancer risk and increased MD are unclear, although several hypotheses exist. The most simple of these suggests that it is merely due to a larger amount of fibroglandular tissue at risk of malignant transformation. However, this is less able to explain the stronger association with percentage dense area compared to absolute dense area (62). It has also been proposed that tissue stiffness and structural changes of stromal collagen may explain this association better (65). Other proposed mechanisms include hormonal factors and altered genetic profile of dense breast tissue (65). Fibroblasts derived from dense and non-dense breast tissue have demonstrated marked differences in gene expression profiles, with several signalling pathways that are known therapeutic targets in cancer being upregulated in dense tissue (66). Attempts to identify genetic associations which increase MD in the population have identified a few polymorphisms which may contribute (67).

Although increased MD appears to be a well-established risk factor, some controversy exists concerning its clinical utility and the reliability of the evidence that it is a risk factor. There have been claims that there are several weaknesses in studies of MD and breast cancer. This includes a lack of standardised definition of high MD (ranging from 25% in some studies to 75% in others) and inappropriate comparisons of the extremes of MD whilst ignoring the majority, weakening the strength of evidence (68). Other issues include the limited impact incorporation of MD has had on the predictive power of risk prediction models (69, 70). Additionally, using MD as a means to assess breast composition may not be reliable, as much of the result relies on thickness of the breast, position and other confounders such as water content of the breast (68). However, research is underway to identify methods of assessing breast density and volume of dense and non-dense tissue which is more accurate (66). Should this be successful and a standardised approach developed, it may be possible to incorporate breast density as a modest predictor of breast cancer risk, alongside other cumulative risk factors.

1.2.5. Ionising Radiation

Ionising radiation (IR) encompasses all X-rays used in diagnostic radiation and radiotherapy. It is known to cause indirect DNA damage to cells by production of free radicals as well as through direct damage via single-stranded or double-stranded DNA breaks. Single-stranded DNA breaks are, more often than not, successfully repaired by base-excision repair mechanisms, however double-stranded breaks are more complex and require repair by homologous recombination repair (HRR) (71). Cumulative DNA damage leaves the cell vulnerable to mutagenesis and cell death. This is taken advantage of in the treatment of various cancers including breast cancer by radiotherapy, where the increased cellular proliferation rates and defective DNA repair mechanisms make the cancerous cells particularly vulnerable to the damage induced by IR. Radiotherapy for breast cancer is known to reduce the risk of local recurrence and improve overall survival (72). However, any radiation exposure also affects normal tissue, which is left vulnerable to bystander DNA damage and in itself may result in mutations which drive cancer development (71).

Research into the effects of IR on breast cancer risks have focussed on specific groups of patients, including those exposed to radiation from nuclear weapons, those treated with IR for Hodgkin's lymphoma, people with scoliosis and tuberculosis monitored with X-rays, and people treated with IR for haemangiomas in childhood. A study into solid cancer incidence was carried out looking at 65,525 female members of a cohort who were resident in Hiroshima and Nagasaki and survived nuclear weapon attacks. The group was stratified by distance from the site of impact at the time of bombing. Excess RR for women over 50 years was 5.3 (95% CI, 2.5-8.6) per Gy unit of exposure (73). Studies of women treated with IR for Hodgkin's lymphoma have shown an overall RR of breast cancer in this group of 5.0 (95% CI, 4.5-5.5), with a tendency to increased risk at younger age of exposure (74). Studies into patients with scoliosis and those treated for haemangioma have shown similar patterns, and risk has been shown to be linearly correlated with dose – there has however been no dose threshold identified, suggesting that any IR exposure may increase risk to some degree (71, 75, 76).

Perhaps of greater significance to the general population who develop breast cancer is the risk of contralateral breast cancer (CBC) after treatment of a primary breast cancer with IR. The evidence for whether or not treatment with radiotherapy increases risk of a subsequent breast cancer is mixed. One study of 134,501 patients with locally invasive or intraductal breast cancer reported a documented CBC in 4.2%. Overall, the risk of CBC was not significantly associated with radiotherapy (RR=1.04, 95%CI, 0.97-1.1) until 5 years after exposure, when radiotherapy was associated with a 14% increase in CBC risk (RR=1.14, 95% CI=1.03-1.26), most pronounced in patients <45 years at initial cancer diagnosis (77)**.** Although inconclusive, this study suggested a trend towards increased risk of a subsequent cancer with a greater effect at younger age of exposure, although other studies of a similar nature have also yielded inconclusive evidence (71)**.** It is plausible that CBCs occurring in those with radiation exposure are just a manifestation of increased genetic risk and other risk factors. In addition, the benefits of radiotherapy at reducing recurrence risk of the breast cancer being treated, probably outweigh the radiation exposure risk; a large meta-analysis of 10,801 women identified an absolute risk reduction of 15.7% for 10-year recurrence (*p*=<0.0001) (78).

Another group of interest is that of *BRCA* mutation carriers whose substantially increased breast cancer risk is due to a genetic defect in proteins required for DNA double-strand break repair by HRR, which may lead to repair by non-homologous end joining (NHEJ), a less reliable method of repair. Since IR is known to induce DNA double-strand breaks, it is reasonable to hypothesize that this group may be more susceptible to the deleterious effects of IR. This is especially relevant since mammography is a suggested means of screening women with known *BRCA* mutations (79)**.** One of the most comprehensive studies into IR exposure and *BRCA* mutation carriers was a retrospective cohort study of 1,993 mutation carriers. They found that any exposure to radiation before the age of 30 was associated with an increased risk of breast cancer (RR=1.90, 95% CI, 1.2-3.0), with risk increasing with dose (80)**.** Levels of dose which had an effect on risk in this cohort was much lower than has been reported in non-*BRCA* mutation carriers, suggesting an increased susceptibility.

Overall, IR is a risk factor for developing breast cancer, with reported RR as high as 4.5 for women exposed to high dose radiation at a young age (71)**.** In specific populations exposed to IR, it may be appropriate for clinicians to have a high index of suspicion for breast cancer. For patients who develop sporadic breast cancer and are treated with radiotherapy at an older age, the risk of developing another cancer attributable to therapeutic IR radiation exposure is probably negligible, however the significance in younger women and *BRCA* mutation carriers is less clear.

1.3. Familial predisposition to breast cancer

Although breast cancer is a common sporadic cancer, up to 20% of cases of breast cancer arise in familial clustering (81, 82), being up to twice as common in a woman with an affected FDR (83). The relationship between FH and breast cancer has long been recognised, despite all the factors contributing to this being unclear. Shared environmental factors are thought to influence risk, however genetic predisposition is thought to be the major factor (84). Some high and intermediate risk genes have been identified, however the mechanisms behind clustering in many families remains unexplained. The genetic risk factors for breast cancer, including known risk genes and FH, are discussed below.

1.3.1. High-risk, high-penetrance genes

Four genes have been identified in which mutations are of high-penetrance and convey a greatly increased risk of breast cancer – these are *BRCA1, BRCA2, TP53* and *PTEN.*

BRCA1 and BRCA2

BRCA1 was first mapped to chromosome 17q21 in 1990 by linkage analysis of 23 families with younger age at cancer onset, bilateral disease and male breast cancer (85). The gene itself was later identified by positional cloning (86). The same techniques led to the identification of *BRCA2* thereafter (87), allowing investigation of both genes roles in development of breast and ovarian cancer. Both *BRCA1* and *BRCA2* are known to be tumour suppressor genes which facilitate the repair of double-stranded DNA breaks (83). DNA double-strand breaks can be repaired by NHEJ or HRR. NHEJ involves the formation of large multiprotein complexes at the site of double-strand breaks and repair of the broken ends via DNA ligases. This is thought to be an error-prone means of DNA repair, however recently a more precise sub-pathway of NHEJ has been identified (88). HRR, or more specifically, gene conversion, involves invasion of the damaged DNA by a strand from the homologous chromosome to act as a template for repair mechanisms. This mechanism is dependent on the DNA recombinase enzyme RAD51, which facilitates invasion of the homologous sequence. Both BRCA1 and BRCA2 proteins are known to aid double-

strand break repair by RAD51 dependent mechanisms through interaction with the RAD51 protein (89, 90). Additionally, both BRCA1 and BRCA2 have been found to be involved in other DNA repair complexes, including BRCA1-Containing Complex, and have roles in cell cycle progression and checkpoint control (91). These functions may be reflected in the genetic instability of tumours with *BRCA* abnormalities, explaining why mutations in the *BRCA* genes increase the probability of carcinogenesis (92).

Disease-causing mutations in the *BRCA* genes have been found to result in the formation of truncated proteins. Most of these are individually rare, however there are known founder mutations in Ashkenazi Jewish and Icelandic populations (83) amongst others. Interestingly, sites of mutation have been found to be linked to the phenotype expressed; mutations in the central portion of both genes have been found to increase the ratio of ovarian to breast cancers within families compared to mutations at the 5' or 3' end (83). The prevalence of *BRCA1* mutations in the UK population is estimated to be at around 0.07-0.09% and for *BRCA2,* 0.14-0.22%, meaning that overall, they contribute to a small fraction of the total breast cancer burden, given that it is such a common disease (93). However women who do carry a *BRCA1* mutation have a 65% and 39% risk by age 70 of breast and ovarian cancer respectively. The corresponding risk for *BRCA2* mutation carriers is 45% and 11% (94). *BRCA1* tumours are more likely to be high grade, ductal carcinomas with a TNBC receptor profile (95). Men who carry a *BRCA* mutation are also at an increased risk of breast cancer, and *BRCA* mutations have been linked to both pancreatic and prostate cancer. Given their role as tumour suppressors, it is unsurprising that homozygosity for mutations in either of the *BRCA* genes is detrimental. There are no reports of homozygous *BRCA1* mutations and it presumed to be embryonically lethal. Those homozygous for a *BRCA2* are known to have Fanconi anaemia, a highly penetrant cancer predisposition syndrome and cause of significant physical abnormalities (83).

TP53

TP53 encodes a transcription factor (TF), p53, which becomes activated in response to cellular stress such as DNA damage, oncogene activation, and hypoxia. Its activation results in downstream activation of tumour suppressor genes involved in DNA repair and cell cycle control. MDM2, an inhibitor of p53 activity is negatively regulated by cellular stress, and the balance between activity of these proteins helps to regulate cellular damage (96). Ultimately, p53 can induce cell cycle arrest by various mechanisms, including at the G1/2 checkpoint by promoting activity of cyclin dependent kinase inhibitor p21 (97). Mutations in *TP53* that disrupt it's function can be found in up to 50% of human cancers (96). However, as well as loss of function, mutations in p53 can result in dominant-negative and gain of function effects (97). Li Fraumeni syndrome is a cancer predisposition syndrome resulting from inherited *TP53* mutations. Women with Li Fraumeni syndrome have a greatly increased lifetime risk of cancer, and carcinoma of the breast has been found to be the most frequent cancer in females of this group (98). Breast cancers occurring in women with a germline *TP53* mutation are significantly more likely to be HER2 amplified tumours (83% compared to 19% of controls, p=1.2x10^{-6}) (99). However, Li Fraumeni syndrome is rare, so while the risk to women with the condition is significant, it's contribution to inherited breast cancer is small.

PTEN

Cytogenetic analysis of a variety of human cancers, followed later by more detailed mapping, identified a tumour suppressor gene at chromosome 10q23.3, called *PTEN.* It was found to be mutated in gliomas, prostate cancer and breast cancer (100-102). Early studies identified loss of heterozygosity at the region containing *PTEN* in around 50% of 32 primary invasive breast cancers (102). In addition, a small study of 5 families meeting diagnostic criteria for Cowden disease, an autosomal dominant, multiple benign hamartoma syndrome, identified a number of *PTEN* point mutations in 4 of 5 families (103). As well as benign hamartomas, individuals with Cowden disease can present with developmental delay and malignant tumours (104). The evidence for *PTEN* as a tumour suppressor was strengthened by the discovery that

heterozygous *PTEN* null mice develop a variety of cancers including breast, thyroid, endometrial and prostate cancer (105). The gene product, PTEN, normally acts to dephosphorylate phosphatidylinositol-3,4,5-trisphosphate, which is an upstream activator of AKT; AKT is a driver of cell proliferation and angiogenesis, and is anti-apoptotic. Therefore, dysregulation of AKT has oncogenic effects (106). Mutations which affect the phosphatase activity of *PTEN* are therefore thought to cause de-repression of AKT and contribute to tumour development. However, mutations out-with the catalytic domain of *PTEN,* which disrupt regulatory features rather than function, have been shown to cause far more diverse phenotypes (for example, autistic spectrum disorder and developmental delay with macrocephaly) (100). The understanding of genotype-phenotype correlations in *PTEN* germline mutation carriers is still developing. However, it has been broadly determined that mutations causing loss of function of the catalytic domain are more likely to cause tumour predisposition phenotypes, and mutations which still confer partial function are more closely linked to autism and developmental delay (100). Studies have determined the lifetime breast cancer risk in patients with a germline *PTEN* mutation to be significantly raised. One study of 3,366 female patients, 295 of which had a *PTEN* mutation estimated the breast cancer risk to rise sharply from age 30 years for those with a mutation, with an approximately 35% penetrance by age 50 years and 85.2% (71.4-99.1%) penetrance by age 70 years. Another study of 154 patients found similar results, with breast cancer risk rising sharply after age 30 years and reaching a 77% (59-91%) risk by age 70 years. Although these studies had limited numbers of patients with a constitutional *PTEN* mutation, they both demonstrate a significantly increased lifetime risk of breast cancer for these patients. Clinically in the UK, women with a *PTEN* mutation are treated as high risk and offered screening at an early age (79). Despite this, *PTEN* mutations are thought to contribute to <1% of breast cancer cases (107).

Although a mutation in one of the highly penetrant genes conveys a significantly increased risk of breast cancer for that individual, mutations in *BRCA1* and *BRCA2* are known to contribute to just 15-20% of the excess familial risk of breast cancer (83). The remaining risk has yet to be explained by shared environmental factors and this explanation for familial clustering has been deemed unlikely. Extensive population studies have added weight to a polygenic model of multiple genetic factors acting independently. A low-penetrance, polygenic risk model has been found to be the best fit for data on women who have had breast cancer, and for multiple-case families not due to *BRCA* mutations. It is unclear whether this polygenic effect would be multiplicative or additive (108). This has been widely accepted as an explanation for the remaining familial breast cancer risk (83). Efforts have therefore been made to identify risk alleles that are of lower penetrance than the genes described above.

1.3.2. Intermediate-risk genes

Several intermediate risk genes have been identified which each confer approximately a 2 fold risk of breast cancer to those with a constitutional mutations in these genes. Some of the genes which have been identified as having a convincing role in breast cancer risk work in DNA repair pathways. *CHEK2, ATM, BRIP1* and *RAD50* are intermediate risk genes which code for proteins which work alongside BRCA1. The RR/OR found to be associated with the intermediate risk genes, discussed below, are shown in *Table 3*. Despite their effect on breast cancer risk, current UK guidance does not suggest screening for mutations in these genes in women with a FH of breast cancer.

CHEK2 and ATM

Among other roles, CHEK2 and ATM are both known to phosphorylate and activate both BRCA1 and p53 in response to DNA damage. CHEK2 itself is also a direct substrate of ATM (109, 110)**.** In activating these proteins, they contribute to BRCA1-mediated double strand break repair as well as cell cycle arrest. A biallelic mutation in *ATM* is known to cause ataxia-telangiectasia, an autosomal recessive condition which predisposes to cancer as well as conferring phenotypic features such as ataxia, telangiectasia and neurodegeneration (110)**.** In a study of individuals with familial breast cancer not found to be caused by mutations in *BRCA1* or *BRCA2,* a heterozygous mutation, *CHEK2**1100delC, was found in 5.1% of breast cancer cases compared to a frequency of 1.1% in healthy individuals (p=3x10^{-7}) (109). Observation of an increased breast cancer incidence in family members of those with ataxia-telangiectasia lead to a similar study looking at *ATM,* which identified mutations in 12/443 affected individuals and 2/521 controls (p=0.0047) (111). These mutations were the same as those known to cause ataxia-telangiectasia, but were present as heterozygous mutations.

BRIP1

BRIP1 encodes BRCA1-associated C-terminal helicase, a protein which interacts directly with BRCA1. It has been found to aid BRCA1 localisation to double-strand breaks through its helicase activity, and be involved in maintaining BRCA1 at double-strand breaks after localisation. Consequently, BRIP1 deficient cells have been found to have delayed double-strand break repair (112). Biallelic mutations in *BRIP1* have been found to cause Fanconi anaemia subtype FA-J (113). Monoallelic mutations in *BRIP1* were identified in 9/1212 individuals with breast cancer from *BRCA* mutation negative families and only 2/2081 controls (p=0.0030) in a British cohort (114). However, several studies from outside the UK have failed to identify an association between *BRIP1* mutations and breast cancer (115). Although *BRIP1* is available in some countries for genetic testing related to breast and ovarian cancer, it has been suggested that the breast cancer risk is in fact negligible, and that *BRIP1* should be considered more of an ovarian cancer risk gene (116).

RAD50

RAD50 constitutes part of the MRN complex (MRE11, RAD50, and NBS1). The MRN complex is phosphorylated and activated by ATM, and interacts directly with BRCA1 in a cell-cycle dependent manner. The complex formed aids in the identification of double-strand breaks and downstream signalling to DNA repair mechanisms (117, 118). Identification of a *RAD50**687delT mutation has been identified in a Finnish population, where it was found to be significantly more frequent in individuals affected with breast cancer (8/317) compared to matched controls (6/1000) (p=0.0004) (119). However, as with *BRIP1* the relevance of this is contentious, as *RAD50* mutations associated with breast cancer have not been linked to breast cancer outwith this study population. Indeed, in a study involving a UK cohort, only 1 of 435 breast cancer cases was found to have a constitutional *RAD50* mutation (120).

PALB2

PALB2 encodes a tumour suppressor protein which is recruited to DNA double-strand breaks in a BRCA1-dependent fashion. Although its recruitment is BRCA1-dependent, the most well documented role of PALB2 is as a partner protein of BRCA2, by facilitating BRCA2 recruitment and stabilisation at double-strand breaks, as well as assisting assembly of RAD51 foci (121). Perhaps unsurprisingly, given its partnership with BRCA2, the *PALB2* gene has been found to cause Fanconi anaemia subtype FA-D1 and FA-N when mutated biallelically, which are both associated with a markedly increased incidence of childhood cancers (83, 122, 123). Monoallelic mutations in *PALB2* have been found to increase susceptibility to breast cancer. A large case-control study found *PALB2* mutations in 10/923 cases of familial breast cancer compared to 0/1,084 control (p=0.0004), a significant finding which has been replicated in other populations (124, 125)**.** Estimates of cumulative risk have been carried out, looking at population specific mutations, including studies in Finnish (126) and Australian (127) populations. These studies found a cumulative breast cancer risk to age 70 years of 40% and 91% respectively. This risk would appear strikingly high, however the analysis in the studies were based on 17 *PALB2* mutation carriers in a cohort of 1,918 breast cancer cases in the Finnish study, and 17 mutation carriers in 1,403 in the Australian study. Perhaps with larger cohorts of mutation carriers, these risks could be further refined.

These intermediate risk mutations are present in genes which encode proteins in the same pathways as BRCA1 and BRCA2. This would suggest that the pathways behind *BRCA*-related breast cancers are also aberrant in these families, but with a different underlying mechanism. It is less clear is why these intermediate-risk mutations should be less penetrant than *BRCA* mutations despite affecting the same pathways. However, there is some debate to be had surrounding this assumption. It has been suggested that in women with a strong FH of breast cancer which is not due to a *BRCA* mutation, a mutation in one of these intermediate penetrance genes still carries a high absolute risk (128). Additionally, in a large study of 35,409 women with a breast cancer diagnosis, 9.3% of women carried a pathogenic variant in a known risk gene, of which 51.5% had a mutation in a gene other than *BRCA1* or *BRCA2* (129). Mutation prevalence in the whole cohort was 2.3% for each *BRCA* gene, and 1% each for *ATM, CHEK2* and *PALB2.* The evidence is now reasonably strong that these 'intermediate-risk' mutations contribute to breast cancer. However, the clinical interpretation of them is less clear. Despite their link with increased breast cancer risk, these genes may account for only 2.3% of familial breast cancer risk (83), and do not explain the entirety of familial breast cancer which is not *BRCA* related, suggesting there are other factors at play.

Table 3 Intermediate risk genes and their associated RR/OR

	Relative Risk (95% CI)	**Odds ratio (95% CI)**	**Reference**
CHEK2	Females 1.7 (1.32-2.20) Males 10.28 (3.54-29.87)	2.52 (0.78-8.18)	The CHEK2-breast cancer consortium (109)
ATM	2.37 (1.51-3.78)	-	Renwick et al. (111)
BRIP1	2.0 (1.2-3.2)	-	Seal S (114)
RAD50	-	4.2 (1.5-12.5)	Heikkinen (119)
PALB2	2.3 (1.4-3.9)		Rahman N (124)

1.3.3. Common, low-penetrance genetic variants

It is thought that genetic risk variants which are common in the population and have a low-penetrance could be responsible for a remaining 28% of familial breast cancer (130). Through genome-wide association studies (GWAS), approximately 76 loci for common, low-penetrance risk variants have been identified using single nucleotide polymorphism (SNP) biomarkers (131). One of the largest studies of this sort in familial breast cancer was carried out by the Breast Cancer Association Consortium (BCAC), which identified 41 new breast cancer associated SNPs, and confirmed the association of 27 previously identified SNPs, having studied women from a widespread, mainly European population (132). The SNPs identified individually carried between a 1.04-1.40 fold increased risk of breast cancer, with an estimated minor allele frequency (MAF) of 10-50%. Overall, they analysed 211,155 SNPs, which were selected based on identification in meta-analysis of previous GWAS, need for fine-mapping of known susceptibility loci and SNPs related to other cancers, among others. 52,675 breast cancer cases and 49,436 controls were then genotyped for these SNPs. Some of the SNPs identified appeared to contribute to breast cancer risk as a whole, however some were found to be specific to sub-types, such as ER +ve or –ve breast cancer. Other studies have identified risk variants which seem only to modify risk in *BRCA1* and *BRCA2* related breast cancers (130). Described below are some of the SNPs identified by the BCAC study which have also been identified in other studies, and for which there is a theorised biological basis for pathogenicity.

A locus identified by the BCAC study, which had been previously reported, was SNP rs2981579 (133), which is in linkage disequilibrium (LD) with rs2981582 (134) identified by Easton et al. (135). Easton et al carried out a 3-stage GWAS for 266,722 selected SNPs. In the first stage, these were genotyped in 390 cases selected for breast cancer FH (at least 2 female FDRs affected) and 364 controls. Of these, 12,711 showed a degree of significance. For the second stage of their study, they aimed to validate these SNPs in 3,990 invasive breast cancer cases and 3,916 controls. For the third stage, 30 of the most significant SNPs were investigated in 21,860 invasive breast cancer cases, 988 breast carcinomas *in situ* and 22,578 controls. 6 SNPs were found to be significant after stage 3, one of which was

rs2981582. They estimated it to have a MAF of 0.3 and a per-allele OR for breast cancer of 1.26 ($p=5x10^{-62}$ at stage 3). An interesting step that was taken in this study was the further analysis of haplotypes associated with the variant using HapMap (136). This confirmed that multiple haplotypes carrying the rs2981582 minor allele were associated with an increased risk of breast cancer, suggesting that the SNP itself, or another variant very strongly associated with it was the driver (135). The SNP lies within intron 2 of the *fibroblast growth factor receptor 2 gene (FGFR2),* a highly conserved region of the genome which contains TF binding sites (135). The FGFR protein family are transmembrane tyrosine kinase receptors which have a role in cell proliferation, survival and apoptosis pathways. Amplification of one of the *FGFR* genes has been identified in 7.5-17% of all breast cancers (137). SNPs in intron 2 have been found to be more common in ER +ve breast cancers, and *FGFR2* has been found to be preferentially amplified in TNBC, with increased expression associated with poorer overall and disease-free survival (137). It was hypothesized that SNPs in TF binding sites could lead to increased binding and protein expression, although some studies have shown there to be a lack of correlation between genotype and cytoplasmic or nuclear protein levels (138). Nonetheless, it is reasonable to speculate that polymorphisms in intronic regions could affect the function or stability of the protein.

Another SNP found to be significant in the BCAC study was rs10069690 (139), which lies within the *telomerase reverse transcriptase (TERT)* gene and carries a RR of 1.18 (95% CI, 1.13-1.25) (132). It had previously been found to be associated with ER -ve breast cancer in an African American population, with a per-allele OR of 1.19 (1.05-1.36, p=0.007) (140). Functionally, *TERT* mutations are well-established in cancer development. The *TERT* gene encodes a rate-limiting catalytic subunit of telomerase which has a role in maintaining genomic integrity through telomere elongation (141). *TERT* it has been found to be overexpressed in several cancers due to promoter mutations, thought to be explained by the immortality associated with telomerase deregulation, which occurs in over 90% of human cancers (141). This often occurs as a somatic mutation in individuals with a normal germline *TERT* gene. However, highly penetrant germline mutations have been identified in familial

melanoma skin cancer (142). Therefore it is possible that less penetrant germline *TERT* mutations in LD or caused by rs10069690 may contribute to familial breast cancer.

Although there is reasonable associative and functional evidence to support a number of SNP associations with breast cancer, the relevance of others identified by GWAS is less clear. For example, rs1045485 (143) aka *CASP8* D302H was only found to be of borderline significance in the BCAC study (132), despite having previously been identified as associated (135, 144). The *caspase 8 (CASP8)* gene is an important initiator of apoptosis, and interestingly, the minor allele identified was deemed to carry decreased risk of breast cancer (per allele OR=0.88 (0.84-0.92) (144). Importantly, there was no association of this with FH, meaning its effect may only be significant in association with specific other polymorphic loci with multiplicative effects. Its role in familial breast cancer is therefore doubtful.

GWAS studies are typically case-control studies which look to identify common variants associated with disease in a population. They are preferred for identifying variants with a high MAF and low penetrance. Although a study population from a multi-centre study such as in the BCAC (132) should provide adequate power, there are several acknowledged limitations of GWAS. The study population for the BCAC study was of mainly European descent, which - although representative of a wide population - means that distant ancestral variation is potentially common to the population. Spurious associations may result given the nature of breast cancer as a common malignancy. Replication of the association in distinct populations is therefore required for reasonable confirmation of the variant with the disease. Several of the 27 previously identified SNPs confirmed in this study had been previously reported in African American and Asian populations, however the newly identified variants would require replication. Another limitation of this study arises from the fact that the cases were unselected for FH of breast cancer. Had women with a significant FH specifically been studied, it may be more confidently speculated that variant associations were giving rise to a familial risk of breast cancer. This is as opposed to being spurious, or due to *de novo* mutations, which is important when considering heritable risk. SNPs are used as genomic

markers in these studies and it is often speculated that when a SNP is identified as associated with disease, either the SNP itself is pathogenic (less likely), or it is lying in LD with a pathogenic variant nearby. Emphasis can therefore be put on fine mapping genes within the region that the SNP is carried. However, many SNPs identified by GWAS lie in regulatory elements and possibly, *trans*-regulatory elements which act from a remote location to the gene they regulate. Some associated SNPs lie within known genes such as *BRCA2* and the *MYC* proto-oncogene (131), however establishing the underlying pathogenic mechanism attributable to these variants may prove challenging without prior mapping and knowledge of the region in which they arise or, alternatively, functional studies.

The impact of such variants on breast cancer risk remains controversial. Estimates for the RR attributable to carrying risk-associated SNPs vary from 33-40 risk alleles required for 3-fold increased risk of breast cancer (130), to 14 risk alleles for a 6-fold increase risk (145). This is dependent on the RR of each SNP, as well as the possibility of allelic interactions. However, identification of variants, and multi-population studies of the risk associated with these variations, is likely to improve understanding of how they contribute to disease. Without adequate understanding of this, it would be challenging to use them clinically. Pharoah et al. suggest that low-penetrance, common variants may be of use for risk-stratification in population programmes (145). Though individually less useful, they state that a small number of susceptibility alleles may be able to discriminate different risk categories. Using absolute risk based on 7 known breast cancer susceptibility loci (RR range from 1.07-1.26), they calculated that women in the 95th risk centile by genotyping would have a 10-year risk of 2.3%, which in the UK would not meet criteria for increased screening (79). However, the top 0.1% would reach a 10-year breast cancer risk of 3%, making them eligible for increased breast cancer screening by mammography. High risk interventions (10-year risk >8%) would therefore, theoretically, be reserved for those most at risk, including *BRCA* mutation carriers. However, it would be reasonable to question the cost-effectiveness of this given the potential need to genotype the many women who attend breast cancer clinics with 2 or more FDRs or second degree relatives (SDRs) diagnosed with breast cancer. These women are

currently considered at 3-8% 10-year risk according to UK guidelines (79). Additionally, incorporation of low-penetrance variant information to the Gail model (146) for breast cancer risk prediction yielded a minimal improvement (130). Despite these reservations, one UK based study assessed a model based on 18 SNPs to reclassify risk for non *BRCA* mutation carriers, and appear to have had some success, with an observed risk 96% of expected according to their model, which assumed a multiplicative effect of published ORs for each SNP (147).

The clinical utility of mapping and genotyping common risk variants is still controversial. However, the contribution of highly-penetrant mutations to familial breast cancer is relatively low, and a proportion of the remaining risk is thought to be explained by intermediate and low penetrance genetic variants. Therefore, it is widely accepted that polygenic risk is a significant contributor to the patterns of breast cancer inheritance seen in some families. With current knowledge, examining patterns of family history and inheritance in these 'non-single gene' cases, and the associated RR of breast cancer, is perhaps more clinically useful than genetic testing. This is an approach that has been taken in UK genetics services (79).

1.3.4. Family history and relative risk of breast cancer

Despite several genes linked to an increased risk of breast cancer, more than 70% of the genetic predisposition remains unaccounted for (83). It has been suggested that the difficulty in identifying predisposing genes may be due to the effect of individually rare variants (perhaps 'family-specific') contributing to increased risk of a common disease. In this case, large population studies may lead to increased difficulty in identifying the causative factors (148). With the lack of knowledge of what causes such a large proportion of breast cancers in women with a FH, risk assessment cannot rely on something as clear cut as genetic testing.

The RR of developing breast cancer is known to be associated with the type, number and age of relatives affected, and so risk calculation can provide a guide as to a woman's risk of developing cancer. A large meta-analysis of 74 studies, which included a variety of ethnic populations and age ranges, and looked at the RR associated with different breast cancer family histories, was carried out by Pharoah et al. in 1997 (149). Using data from 52 case-control and 22 cohort studies, they calculated the pooled RR of breast cancer for various family histories. These are summarised in *Table 4*. They also acknowledged that the majority of studies saw the RR increase if the relative was diagnosed at <50 years and also in subjects <50 years.

Another large meta-analysis carried out by the Collaborative Group on Hormonal Risk Factors in Breast Cancer also estimated the RR of various family histories. They found that the RR of breast cancer increased with an increasing number of affected FDRs ($p<0.001$) with a RR for 1, 2 and 3 affected FDRs of 1.8 (95% CI, 1.7-1.91), 2.93 (2.37-3.63) and 3.9 (2.03-7.48) respectively (150). Although they did not calculate RR for any history of affected FDR and so are not directly comparable with Pharaoh et al. the results of their analysis are broadly similar, suggesting an approximate two-fold risk of breast cancer in women with at least 1 affected FDR, as well as an increase in risk for women with relatives diagnosed at <50 years (150).

The Nurses' Health Study, a prospective cohort study of 121,701 US nurses also sought to establish the RR associated with FH, and by age classification of the

relative at diagnosis (151). They appeared to find slightly lower RR associated with a first degree FH than the previously mentioned studies, identifying a RR of only 1.69 (95% CI, 1.39-2.05) and 1.37 (1.22-1.53) for a mother diagnosed at <50 years and >50 years respectively. The corresponding RR for a sister were 1.66 (1.38-1.99) and 1.52 (1.29-1.77). Although they found a trend towards increased risk for women with a FDR diagnosed <50 years, this was not statistically significant (151). They suggest that the reason for the difference between their results and that of meta-analyses, is that as they undertook a prospective cohort study therefore recall bias which may be present in retrospective cohort studies used in meta-analyses is eliminated.

Although the meta-analyses by Pharoah et al. and the Collaborative Group on Hormonal Risk Factors in Breast Cancer carry great statistical power, meta-analyses are subject to a great deal of publication bias which may influence results. Despite this, it is particularly striking that Pharoah et al. found that having any affected FDR carried a RR of 1.9 (1.7-2.0) irrespective of age of the relative at diagnosis (149), since as many as 15-30% of women have been found to report a FH of breast cancer (151, 152).

Table 4. RR associated with different family histories according to meta-analysis by Pharoah et al. (149)

Relative Affected	**RR (95% CI)**
Any relative (unspecified)	1.9 (1.7-2.0)
FDR (mother, sister or daughter)	2.1 (2.0-2.2)
Mother	2.0 (1.8-2.1)
Sister	2.3 (2.1-2.4)
Daughter	1.8 (1.6-2.0)
Mother and sister	3.6 (2.5-5.0)
SDR	1.5 (1.4-1.6)

Although a FH of breast cancer carries an increased risk of developing the disease, a positive FH has been found to be associated with smaller tumour size. This is thought to be due to a higher degree of vigilance and screening uptake in the population with a FH. They are therefore are more likely to be diagnosed whilst asymptomatic (152). This demonstrates a potential benefit in asymptomatic screening, as it allows earlier recognition of cancers in women who are at risk, potentially making the tumours more treatable. Since women with a FH are also diagnosed at a younger age (152) than those without, early screening in this population is likely to be of benefit. One study found no statistically significant difference in prognosis between women with and without a FH of breast cancer, however it was unclear whether or not the population studied had been screen-detected or symptomatically detected (153). Mammographic screening in the UK has been found to be effective at decreasing the incidence of invasive cancers in those who have had DCIS screen detected and treated (154). Additionally, the cost effectiveness and incidence of negative effects, such as false positive results, has been found to be reduced in screening programmes which alter screening based on risk (155). Therefore, in order to achieve early detection of breast cancer in higher risk women, whilst avoiding unnecessary cost, identifying an appropriate risk cut off and FH that this corresponds to is desirable.

1.4. Risk Assessment Models

1.4.1. Gail

One of the first widely used models for breast cancer risk prediction using FH, was developed by Gail et al. in 1989 (156). They developed their model using case-control data from the Breast Cancer Detection Demonstration Project, a multi-centre breast cancer screening programme, which enrolled over 280,000 women for annual breast screening over a 5-year period (157). Major risk factors predictive of breast cancer which they identified in their cohort included late age at first childbirth, multiple previous benign breast biopsies, early menarche, and FH of breast cancer in an FDR. Gail et al. assumed a proportional hazards model – the addition of each individual risk factor has a multiplicative effect on the baseline hazard rate of breast cancer. The RR associated with each risk factor was calculated (*Table 5*), allowing an individual's cumulative RR to be determined by multiplication of their RR for each covariate. They then provide a means to convert the individual's RR, into % projected probability of developing breast cancer within the next 10, 20, or 30 years, dependent on their initial age at risk calculation. From their model, Gail et al. found consistent results with previous studies investigating the probability of developing breast cancer in women with a first degree FH (156, 158, 159).

Validation of the Gail model using the Nurses' Health Study (160) cohort found that overall, by comparing expected to observed breast cancer incidence, the Gail model over predicted breast cancer risk in up to 33% of women (161). Similarly, a smaller validation study found that the Gail model tended to over predict incidence, particularly in the younger age group (162). In their original publication, the authors allude to the possibility of overestimation in younger women. They reason that due to their study population undergoing annual mammography, the model best applies to this group, in which frequent screening identifies early cancers which may otherwise not be apparent until a later date. Therefore overestimation may occur in younger unscreened women (146). In agreement with this, validation studies which identified overestimation cited this as a possible cause (161, 162), and suggested that if this is a valid explanation, cancers in less frequently screened women may only be identified at a later disease stage (162). Concern has also been raised that the Gail model is limited in its incorporation of FH. Since it only accounts for first degree FH, it runs the risk of underestimating risk prediction in those with a FH of bilateral

breast cancer, early onset breast cancer, ovarian cancer, a second degree FH etc. A study by Euhus et al. compared the suitability of the Gail model compared to other risk prediction models in a cohort enriched for those with a FH of breast cancer (163). Of their cohort, 74% had a history of risk factors which would be expected to be problematic for the Gail model. It was thought to underestimate risk in 13%, due to its inability to account for elements of breast cancer FH other than incidence in FDRs (163). However, a systematic review which looked at 16 papers validating the Gail model was inconclusive regarding the calibration of the Gail model, stating that there was significant heterogeneity between papers and that the model did not appear to perform consistently well (164).

Table 5. RR associated with risk factors identified by Gail et al. 1989 (156)

Risk Factor		Associated relative risk*
Age at menarche (years)		
≥14		1.000
12-13		1.099
<12		1.207
Number of previous benign biopsies		
Age <50 years		
0		1.000
1		1.698
≥2		2.882
Age ≥50 years		
0		1.000
1		1.273
≥2		1.620
Age at first live birth (years)**	Number of FDRs with a history of breast cancer**	
<20	0	1.000
	1	2.607
	≥2	6.798
20-24	0	1.244
	1	2.681
	≥2	5.775
25-29 or nulliparous	0	1.548
	1	2.756
	≥2	4.907
≥30	0	1.927
	1	2.834
	≥2	4.169

*As calculated by Gail et al. **RR for age at first live birth and number of FDRs with a history of breast cancer is expressed as a combination of the two factors.

1.4.2. Claus

Another study to investigate risk factors in breast cancer using a proportional hazards model was carried out by Claus et al, using 4,730 breast cancer cases and 4,688 controls from the Cancer and Steroid Hormone Study (165). However, unlike the analysis by Gail et al., it focussed on identifying risk factors relevant to women at a possible increased risk of breast cancer due to their FH. The risk factors investigated were bilaterality, history of benign breast disease, parity, number of livebirths and stillbirths, age at first pregnancy, age at onset of breast cancer, age at menarche, menopausal status, and age at last menstrual period. Whilst this information was gathered on cases and controls, it was incorporated into analysis of risk of breast cancer for mothers and sisters of the cases/controls. Also analysed was the age at onset of breast cancer for affected relatives. Breast cancer cases were then divided into groups, according to the number and type of relatives affected with breast cancer. The distribution of risk factors was assessed across varying FH patterns, to identify those which could be linked to a genetic predisposition to breast cancer. The results of their investigations showed that relatives of cases were more likely to go on to develop breast cancer than relatives of controls, with risk increasing as age at onset of the case decreased, and with increasing number of affected relatives. However, the distribution of the other risk factors investigated showed no relation with FH, suggested that they had a less pronounced impact on breast cancer risk in families with a possible genetic predisposition (165). Further analysis by goodness-of-fit tests, to compare the observed age-specific risk with predicted risk under genetic models was carried out. They concluded that most of the incidence of breast cancer was likely to be due to non-genetic factors. However, a small subset of disease may be accounted for by an autosomal dominant allele segregating for increased susceptibility to breast cancer, particularly early-onset, for which they proposed *BRCA1* as a possible candidate (166, 167). With the discovery of new high and intermediate risk genes, this single gene dominant model is largely disputed, however, this was the premise upon which the Claus model was developed.

Claus et al. proceeded to use data from their analyses to develop a risk prediction tool for clinicians assessing breast cancer risk in a woman with a FH (167).

Using the type and number of affected relatives, and age at onset, they produced tables which could be used to estimate the age-specific, cumulative probability of developing breast cancer for women with a FH. The tables are classified according to breast cancer FH, including one FDR affected; one SDR affected; two FDRs affected; mother and maternal aunt affected; mother and paternal aunt affected; one maternal and one paternal SDR; and two SDRs (both maternal or both paternal) affected. All tables are stratified by age of the individual and affected relatives (167). In a separate publication, they also evaluated the impact having a first degree FH of ovarian cancer had on breast cancer risk under the assumption that transmission of both cancers could partly be explained by the same rare autosomal dominant allele. As a result, they produced similar cumulative probability tables according to FH. They found that a woman with a first-degree FH of ovarian cancer had a 50% greater chance of developing breast cancer than a woman with no FH of ovarian cancer (168).

As with other models, the Claus tables are subject to limitations. They do not take into account information on non-hereditary risk factors, and can only accommodate for a FH of a maximum of two affected relatives, meaning they may underestimate risk in families with a bigger number of affected relatives. Additionally, they are based on risk estimates for women in the USA in the 1980s, however the incidence rates in North America and in Europe have varied from the cohort the tables were based on since their development (169). Indeed, Amir et al. found the original data gathered by Claus et al. (on which the tables are based) to underestimate breast cancer risk, with an expected to observed (E/O) ratio of 0.56 in a screened population with a FH of breast cancer (170). Additionally, another study comparing risk models found the Claus tables to estimate a lower risk of breast cancer compared to other prediction models. Despite this study comparing risk models in a very small number of patients, they concluded that the Claus model is likely to underestimate breast cancer risk (171). An extended version of the Claus tables, ‘Claus-Extended’ has been developed to incorporate ovarian cancer, bilateral breast cancer and multiple cases of breast cancer (172), and the risk estimates obtained from this have been found to be more in agreement with more modern risk assessment models (171).

1.4.3. BRCAPRO

With the increase in understanding of the importance of *BRCA1* and *BRCA2* in breast cancer pathogenesis, and the identification of the genes themselves, testing women for these genes became a possibility in order to identify those at particularly high risk (86, 87, 173). In 1998, Parmigiani et al. developed a method for determining the probability that an individual carries a germ-line mutation in *BRCA1* or *BRCA2* on the basis of their FH, assuming an autosomal dominant mode of inheritance (174). In their model, they use Bayesian methodology to incorporate first and second-degree FH of breast and ovarian cancer into a calculation of probability. The prior probability of carrying a mutation in either *BRCA1* or *BRCA2* is determined using published mutation frequency and prevalence rates (175, 176). The estimated true probability of the individual being a carrier is then inferred. This is done by collecting information concerning each family member - whether they have been diagnosed with breast or ovarian cancer, the age at diagnosis, or, if cancer free, their current age or age at death - using published data by Shattuck-Eidens et al (177). Additionally, they incorporate the penetrance of *BRCA1* and *BRCA2,* and allow for uncertainty in carrier rates and penetrance functions (174). This model was subsequently incorporated into the *BRCA1/BRCA2* carrier probability tool 'BRCAPRO' (178).

To assess the validity of BRCAPRO, a retrospective analysis of families in which at least one family member had been tested for both *BRCA1* and *BRCA2* was carried out by Berry et al. (179). They concluded BRCAPRO to be a highly sensitive counselling tool, with a mean carrier probability of 79% for those who test positive and 39% for those who test negative (p=10^{-10}). However, the tool still missed approximately 15% of mutations (179). A similar value of 16% of mutations were missed by BRCAPRO in a study which used a carrier probability of 10% as a threshold for testing (180). This same study compared carrier probability prediction by BRCAPRO to the subjective estimates of risk counsellors. It concluded that while BRCAPRO was more specific and better able to discriminate between mutation carriers and non-carriers, input from risk counsellors was still required due to BRCAPROs incorrect negative assignment of mutation-carrying families (180). Another study found that while BRCAPRO performed well in ranking individuals by carrier probability, it may over predict

mutations, particularly in *BRCA1 (181)*. Some of the inaccuracy of the original BRCAPRO model may be due to discrepancies in the published data relied upon to build it. Another drawback of the original model is that the authors chose to consider *BRCA1* and *BRCA2* and regard all other breast cancer as being sporadic. However, it is accepted that there are families for whom an increased risk of breast cancer is apparent, but the risk is not due to *BRCA1* or *BRCA2*. Whilst the purpose of BRCAPRO is to estimate carrier probability of mutation in these genes, it is unable to estimate the risk of a woman with a FH developing breast cancer, despite carrier status.

Since its original development, BRCAPRO has undergone several updates. It was updated to account for the impact that medical interventions may have on carrier probability, specifically the impact of mastectomy and oophorectomy (182, 183), which minimise the risk of cancer in *BRCA* mutation carriers (184, 185). This was thought to improve accuracy versus other carrier probability prediction models that did not account for medical interventions within the family (182). Additionally, breast tumour markers were added to the prediction model. Breast cancer in *BRCA1* carriers are more likely to be TNBCs, whereas *BRCA2* carriers and non-carriers are more likely to have ER/PR positive breast cancer. HER2 receptor status as been found to be similar between *BRCA1* and non-carrier breast cancer (186). Addition of these markers to BRCAPRO was found to improve discrimination between carriers and non-carriers as well as between *BRCA1* and *BRCA2* (187). Various other improvements to the programme have been made, including incorporating ethnicity, CBC risk, and the ability to account for missing details such as age, with each addition claiming to further improve the tools predictive ability (183, 188). BRCAPRO has been incorporated into the BayesMendel R statistical package, designed to calculate risk prediction for different inherited cancers according to FH. As well as providing information on carrier probability, this package also gives as an output the future risk of breast and ovarian cancer for individuals (183, 189).

1.4.4. BOADICEA

An analysis on a population derived from the Anglican Breast Cancer Study Group (93) with breast cancer onset before age 55 was carried out by Antoniou et al., to investigate different genetic models for breast and ovarian cancer, and their use as a risk prediction model (190). From the 1,484 breast cancer cases, information was gathered regarding FH, and DNA was extracted for analysis of *BRCA1* and *BRCA2* mutations. In addition, another group of families were recruited; these contained two or more breast cancer cases with one diagnosed under age 50, and at least one family member tested for *BRCA1* or *BRCA2.* Using the combined data from these groups, the analysis modelled the effects of *BRCA1* and *BRCA2* alongside either a hypothetical major breast cancer susceptibility locus *'BRCA3'* or a polygenic effect. The model of best fit for determining breast cancer susceptibility was found to be the *BRCA1*/*BRCA2* mutation, plus a multiplicative polygenetic effect which has the same effect regardless of *BRCA1/2* mutation status (83).

They found the risk of breast cancer for non-carriers over all polygenic effects to be 4.89%, however by percentile distribution, the lowest and highest percentiles had a 1.3% and 17.9% risk respectively. Taking into account superimposed polygenic effects, they estimated the cumulative breast cancer risk by age 70 for a *BRCA1* mutation carrier to be 35.26%, and for a *BRCA2* mutation carrier 50.26%. On analysis of breast cancer risk in smaller age categories they found that the risk of breast cancer with a *BRCA1* mutation increases till 40-49 years of age and then decreases. However, the risk of breast cancer for *BRCA2* mutation carriers was found to continually increase with age. For ovarian cancer, the risk by age 70 was 26% for *BRCA1* carriers and 9.1% for *BRCA2* carriers, with non-mutation carriers having a 1% probability. No polygenic component was applied to ovarian cancer risk prediction (190).

Using the results of their analysis, a model named 'Breast and Ovarian Analysis of Disease Incidence and Carrier Estimation Algorithm' (BOADICEA) was developed to determine mutation carrier probabilities and cancer risks based on FH of breast cancer; the model was then validated using the results of epidemiological studies (191) and later implemented into a web-based software programme. In

agreement with previous studies, BOADICEA shows increasing risk of breast cancer, and mutation carrier probability with an increasing number of relatives affected and decreasing age at diagnosis. To compare BOADICEA to empirical data, the age specific risk of breast cancer for a woman whose mother was affected at the same age was calculated. On comparison of the familial RR prediction made by BOADICEA with observed values from the meta-analysis by the Collaborative Group in Hormonal Factors in Breast Cancer (150), BOADICEA was found to have a correlation coefficient of 0.99. Another analysis however did not produce as strong a correlation; the age specific annual incidence, according to BOADICEA, in a sister of a breast cancer patient was calculated, and compared to an analysis by Peto and Mack (192). The correlation coefficient under these circumstances was found only to be 0.36. However, in both of these analyses, BOADICEA was found to have as good as, or better correlation with observed data than the Claus model or the original BRCAPRO model (*Table 6*) (191). Another study found BOADICEA to outperform BRCAPRO when looking at an Ashkenazi Jewish population, whereas the previously mentioned study found BRCAPRO to perform marginally better than BOADICEA in a non-Ashkenazi Jewish population due to it being slightly better at predicting *BRCA1/2* mutations (193).

Table 6. Correlation coefficients of BOADICEA, BRCAPRO and Claus

	Intraclass correlation coefficient between predicted and observed data[a]	
	Age specific risk for a woman with a mother affected at her current age[b]	Age specific annual incidence for a sister of a women affected with breast cancer[c]
BOADICEA	0.99	0.36
BRCAPRO	0.35	0.07
Claus	0.99	0.34

[a]as reported by Antoniou et al (94) [b]Observed data from the Collaborative Group on Hormonal Factors in Breast Cancer (150). [c]Observed data from Peto and Mack (192)

BOADICEA has also been validated in international studies. An evaluation of BOADICEA in women with a FH of breast cancer in a Swedish population found that its ability to discriminate between *BRCA* mutation carriers and non-carriers was measured to an area under the receiver operating curve (AUROC) of 0.83 (194). The ratio of observed: expected (O:E) invasive cancer was 1.41, and the ability of the model to discriminate between those who developed invasive cancer and those who did not had an AUROC of 0.62, slightly better than chance for risk prediction. They concluded that, in a Swedish population, BOADICEA could be viably used as a tool for mutation carrier probability and lifetime breast cancer risk calculation (194). However, BOADIECA performed poorly at predicting the number of *BRCA1* or *BRCA2* mutations specifically. This was thought to be due to a high number of *BRCA1* mutations in Sweden, due to a number of founder mutations. This reflects an inherent problem with all risk prediction models based on population studies, that the model may only be reliable within that population due to genetic, lifestyle and cultural differences across societies. Interestingly however, BOADICEA was found to fit well to data derived from high risk French-Canadian families, suggesting that the allele frequency of *BRCA* mutations in this group is similar to the UK data the model was based on (181).

BOADICEA claims to be advantageous compared to other risk prediction models as it has the ability to include information on all available relatives rather than being limited to FDRs or SDRs, unlike other models (191). The nature of the model means that the addition of other relevant information is possible. It was updated to include the risk of male breast, prostate and pancreatic cancer (195). In addition to this, the breast and ovarian cancer incidences in mutation carriers were updated to be based on a larger number of families, in order to increase reliability. The polygenic component was also altered to be age dependent as opposed to constant across all ages (195). Tumour markers including ER, PR and HER2 status have also been incorporated (196). Independent validation of the updated model found it to be well-calibrated, with an expected to observed number of breast cancers of 0.92 and an AUROC of 0.7 (197). Criticism of the BOADICEA model has implied that, as well as concern that it potentially underestimates the importance of ovarian cancer, the

model may be difficult for use by clinicians and therefore not suited for routine clinical practise (198). In addition, if polygenic mode of inheritance is assumed in a family, inclusion of relatives beyond FDRs and SDRs may be thought to be unnecessary due to the rate of decay in shared genetic factors beyond this. Despite this, BOADICEA is named as a recommended prediction model by the National Institute for Health and Care Excellence (NICE) guidelines on familial breast cancer (CG164) (79).

1.4.5. Tyrer-Cuzick/ 'IBIS'

Using published data from a number of sources, Tyrer, Duffy and Cuzick developed a risk prediction model which incorporates genetic, familial and - like the Gail model - personal risk factors to predict absolute risk of breast cancer (199). They use a two locus, dominant genetic model, with one locus representative of both *BRCA1* and *BRCA2*; the other locus was designed to represent additional unknown 'low penetrance' susceptibility genes which contribute to the higher risk of breast cancer in those with a FH. The probability of a *BRCA* mutation and subsequent breast cancer risk, given FH, was ascertained using data published by Ford et al. (200). The additional risk instated by the presence of the low penetrance susceptibility allele, was estimated using observed data from Anderson et al. via follow up of women with a FH of breast cancer (201). The age specific risk of breast cancer for each genotype was calculated, however in addition to this, the personal risk factors were added into the calculation. This included age at menarche, parity, age at first childbirth, age at menopause, atypical hyperplasia, LCIS, height and BMI. These risk factors were assumed to be multiplicative, and when superimposed on the risk derived from FH, give the final risk estimation from the Tyrer-Cuzick model.

The Tyrer-Cuzick model was validated in a study using data from 1,933 women who had attended for evaluation of their FH of breast cancer, over a mean follow up time of 5.27 years. The AUROC for the Tyrer-Cuzick model was found to be 0.762, with an E/O ratio of 1.09, outperforming both the Claus and Gail model (170). Additionally, as it is a relatively newer model, the baseline risk estimates of breast cancer used are thought to be more accurate than that of the older risk models (171). With regard to *BRCA* mutations, a German study found Tyrer-Cuzick to successfully predict the total number of mutations in their cohort. However, it tended to over predict mutation rate in lower risk families (202). Several publications have found the Tyrer-Cuzick model to be outperformed by BOADICEA and BRCAPRO (203, 204). There has also been evidence to suggest that the model may be poor at discriminating between high and low risk women, and may over predict risk in some groups (205). Whilst it is theoretically beneficial to include both personal and genetic risk factors for breast cancer risk prediction, in clinical practise, the addition of

multiple variables has the potential to lead to problems; should the necessary information (such as that of previous breast biopsies) be unavailable to the clinician, the result may be a less reliable due to missing information, resulting in an artificially low, or high risk estimation (170).

1.4.6. Manchester Scoring System (MSS)

There appears to be a trade-off between the complexity and time-consuming nature of computer-based models vs. the over-simplicity of manual models. To address this, Evans et al. sought to develop a model for *BRCA1/2* carrier prediction that was simple and efficient enough for use in busy clinical practice. At the same time, they aimed to avoid the short comings of other manual models, such as only taking into account a limited number of affected relatives (206). By analysis of the pedigree information, and mutation analysis of 422 families in North West England, they develop the Manchester Scoring System (MSS). This is an empirical scoring system, incorporating age-stratified FH of breast, male breast, ovarian, pancreatic and prostate cancer. Separate scoring was developed for *BRCA1* and *BRCA2* (*Table 7*). A score of >10 was stated to correlate to a 10% risk of finding a *BRCA* mutation in the family. Notably, in their validation of the original model, Evans et al. calculated AUROC statistics for the MSS and other models, and found MSS to outperform each of them, including computer model BRCAPRO (206). The MSS had an AUROC of 0.772 for combined *BRCA1* and *BRCA2* scores, compared to 0.596 for BRCAPRO in their validation group.

Table 7. Original MSS. Adapted from Evans et al. 2004 (206).

Cancer (age at diagnosis)	***BRCA1***	***BRCA2***
FBC[A] (<30)	6	5
FBC (30-39)	4	4
FBC (40-49)	3	3
FBC (50-59)	2	2
FBC (>59)	1	1
MBC[B] (<60)	5 (if *BRCA2* already tested)	8
MBC (>59)	5 (if *BRCA2* already tested)	5
Ovarian cancer (<60)	8	5 (if *BRCA1* already tested)
Ovarian cancer (>59)	5	5 (if *BRCA1* already tested)
Pancreatic cancer (any age)	0	1
Prostate cancer (<60)	0	2
Prostate cancer (>59)	0	1

[A]Female Breast Cancer [B]Male Breast Cancer. Scores are summated for each cancer in a direct lineage; score >10 is equivalent to a 10% chance of identifying a *BRCA1/BRCA2* mutation on testing.

With evidence emerging of the impact of pathology information in relation to *BRCA* mutation status, the MSS was updated in 2009, in an attempt to further improve the sensitivity and specificity of the scoring system. No changes were made to *BRCA2* scoring however information on grade, type and HR status of breast cancer, as well as ovarian cancer pathology was added to the scoring for *BRCA1* (207). Most notably, 4 scoring points were added for *BRCA1* if a grade 3 TNBC was present, and 4 points were to be deducted for *BRCA1* if a HER2 positive breast cancer was present in the family. Despite these changes only being applicable to *BRCA1*, they found that the AUROC for combined *BRCA1/2* score increased from 0.701 to 0.726 at a 10% threshold in the study population used to validate the updated model (207). Despite this improvement, it is important to note that pathology information for cancer in a pedigree may not be readily available to clinicians in cancer genetics services. Indeed, even in creation of the model it was stated that specific information on breast pathology was only available in 43% of cases (207).

Independent validation for both versions of the MSS was carried out in a German population, with results similar to that of the original validation by Evans et al. The combined *BRCA1/2* AUROC for the 2004 MSS was 0.77, with an improvement to 0.80 in the 2009 version. *BRCA1* prediction was found to be significantly improved by use of the 2009 version, where pathology information was available. However, on further analysis of the German cohort, pancreatic cancer was not found to be significantly predictive for either *BRCA1* or *BRCA2*, and the presence of female breast cancer >50 was found to be negatively predictive, in contrast to the MSS (208). Despite this, the authors concluded that the MSS provided a valuable and easy to use tool for identifying possible *BRCA* mutation families, with similar AUROC values to BOADICEA and BRCAPRO in studies of the same cohort (202, 208). In agreement, another study found the difference in AUROC statistics between MSS and BRCAPRO to be insufficiently different to claim superiority of one over the other (209). Interestingly, a study was carried out to assess threshold for MSS to achieve a 90% sensitivity in a Canadian population. They found that a score of 7.58 was necessary to achieve 90% sensitivity, much lower than the suggested score of 10 (203). Additionally, they found that at the conventional threshold, the MSS (AUROC 0.68)

was outperformed by BRCAPRO and BOADICEA (0.76 and 0.74 respectively) in combined *BRCA1/2* identification. However, it performed better than Tyrer-Cuzick (AUROC 0.47). It is worth noting that the study found BOADICEA and BRCAPRO to be the most difficult to use, with the data required often difficult for patients to recall, whereas the MSS was found to be one of the most convenient to use in practice (203).

Notable drawbacks of the MSS include the potential unavailability of pathological parameters, and failure to take into account full pedigree structure. It is also not suitable for use in populations with founder mutations. Nonetheless, alongside BOADICEA, it is suggested as a recommended scoring system for identifying women at risk of *BRCA* mutations in the NICE guidelines for management of familial breast cancer (79). Given its simplicity of use, the MSS is a popular choice in clinical practise. The Institute of Cancer Research (ICR) undertook extensive evaluation of the criteria which would meet the 10% carrier probability threshold suggested by NICE (79) for *BRCA* mutation testing. Based on this, they produced a modified version of the MSS (MMSS) at which a score of 17 or greater warrants testing for *BRCA1* and *BRCA2* mutations (see *Table 8*) (79, 210).

Table 8. ICR MSS. Scoring system is combined for *BRCA1* and *BRCA2*.

Manchester Score	
Cancer (age at diagnosis in years)	Score
FBC[A] (<30)	11
FBC (30-29)	8
FBC (40-49)	6
FBC (50-59)	4
FBC (>59)	2
MBC[B] (<60)	13
MBC (>59)	10
Ovarian cancer (<60)	13
Ovarian cancer (>59)	10
Pancreatic cancer (any age)	1
Prostate cancer (<60)	2
Prostate cancer (>59)	1

[A]Female Breast Cancer. [B]Male Breast Cancer.

1.4.7. National Institute for Health and Care Excellence

NICE, established in 1999, is a UK organisation which develops evidence-based guidance for healthcare. Although health care professionals working in the UK National Health Service (NHS) are not legally obliged to follow the guidance provided by NICE, they are actively encouraged to do so. This makes NICE recommendations an important influence on clinical practice in the NHS. The current NICE guidance on classification and management of people with a FH of breast cancer (CG164) was published in June 2013 (updated August 2015) (79). The guideline stratifies women with a FH of breast cancer, according to their risk, both for women with and without a personal history. Dependent on this, they suggest appropriate surveillance regimes for these women. Recommendations are offered for primary care workers as to the appropriateness of referral to secondary care and specialist genetics services in a women presenting with concerns about her breast cancer risk. The clinical evidence review used to build the guideline is publicly available via the NICE website (211). To address the best method of predicting individual risk of developing breast cancer in a woman with a FH they reviewed literature assessing and comparing risk models including Gail, Claus, Ford, and BRCAPRO. They conclude that existing computer models may underestimate risk through FH, and the degree of correlation between risk models is poor. Further review was carried out to address the optimal method of calculating carrier probability, comparing the sensitivity and specificity of a range of models including BRCAPRO, BOADICEA, Tyrer-Cuzick, IBIS and the MSS. They found consistent evidence that all of the carrier probability prediction models perform better than chance, however BOADICEA was the best calibrated model. It is unsurprising therefore, that CG164 recommends BOADICEA as an acceptable method of carrier probability calculation. In addition to this, the MSS is also recommended, perhaps as the clinical evidence review acknowledges it may be more practical for use by clinicians (211). The threshold for mutation testing is set at greater than 10% risk of being a carrier. Aside from carrier probability, BOADICEA can provide a 10-year, or lifetime risk estimation for developing breast cancer, unlike MSS. NICE stratify women into low/population, moderate and high risk based on percentage absolute lifetime risk and percentage 10-year absolute risk from age 40, as demonstrated in *Table 9*. In the previous NICE guideline for familial breast cancer

(CG041) (212), criteria equivalent to these risk categories were given, based on data from Claus et al. and the Collaborative Group on Hormonal Factors in Breast Cancer (150, 167). These criteria, demonstrated in *Table 10*, are still present in the updated guideline CG164, and are used in clinical practise for risk stratification, as well as having been used in research in this format by Evans et al. (213).

Table 9. Breast cancer risk category by lifetime risk and risk between ages 40-50. (Adapted from NICE guideline CG164 (79))

	Breast cancer risk category		
	Low risk (population risk)	Moderate risk	High risk[A]
Lifetime risk (from age 20)	≤17%	>17% but <30%	≥30%
Risk between ages 40-50	<3%	3-8%	>8%

[A]Including *BRCA1* and *BRCA2* mutation carriers and those with other conditions/mutations known to predispose to breast cancer.

Table 10. Criteria for categorisation into moderate or high risk category by FH alone according to NICE CG164 (79).

NICE moderate risk criteria	- One FDR diagnosed with breast cancer at younger than age 40 years **or** - Two first-degree or SDRs diagnosed with breast cancer at an average age of older than 50 years **or** - Three first-degree or SDRs diagnosed with breast cancer at an average age of older than 60 years
NICE high risk criteria	*At least the following female breast cancers only in the family:* - Two first-degree or SDRs diagnosed with breast cancer at younger than an average age of 50 years (at least one must be a FDR) **or** - Three first-degree or SDRs diagnosed with breast cancer at younger than an average age of 60 years (at least one must be a FDR) **or** - Four relatives diagnosed with breast cancer at any age (at least one must be a FDR) **or** *Families containing one relative with ovarian cancer at any age and, on the same side of the family*: - One FDR (including the relative with ovarian cancer) or SDR diagnosed with breast cancer at younger than age 50 years. **or** - Two first-degree or SDRs diagnosed with breast cancer at younger than an average age of 60 years. **or** - Another ovarian cancer at any age. **or** *Families affected by bilateral cancer (each breast cancer has the same count value as one relative):* - One FDR with cancer diagnosed in both breasts at younger than an average age 50 years. **or** - One first-degree or SDR diagnosed with bilateral cancer and one first or SDR diagnosed with breast cancer at younger than an average age 60 years. **or** *Families containing male breast cancer at any age and, on the same side of the family, at least:* - One first-degree or SDR diagnosed with breast cancer at younger than age 50 years. **or** - Two first-degree or SDRs diagnosed with breast cancer at younger than an average age of 60 years.

NICE recommend additional screening for women at moderate and high risk, as seen in *Table 11*, in the form of mammographic surveillance or MRI. This additional screening is of particular relevance for younger women who are not yet of the age to be enrolled in the NBSP. A study of 62 breast cancers in screened and 1,108 non-screening women <50 years was carried out to compare survival outcomes The screened women had a FH of cancer and were screened at 12-18 month intervals. It demonstrated that cancers picked up on screening were significantly more likely to be smaller and significantly less likely to be invasive or to have metastasized to lymph nodes, despite similarity in cancer grade between the two groups. As a result, there was a significantly lower proportion of recurrences and breast cancer deaths in the screened population (p=0.008) (214), suggesting a benefit in screening younger women with a FH of cancer. A similar but much larger study of 6,710 screened and 106,971 unscreened women <50 years, also reported benefit of mammograms, reporting significantly lower 10-year mortality rates (p=0.022), and in addition, smaller tumours of a more favourable grade, which were less likely to have metastasised to lymph nodes (215).

Table 11. Recommended screening and intervention by risk category according to NICE CG164 (79).

Screening/Intervention	Moderate Risk	High Risk (including *BRCA* mutation carriers)
Mammographic surveillance	*Offer annually to women:* - aged 40–49 years *Consider annually for women:* -aged 50-59 years	*Offer annually to women:* - Aged 40–59 years at high risk of breast cancer but with a 30% or lower probability of being a *BRCA* or *TP53* carrier - Aged 40–59 years who have not had genetic testing but have a greater than 30% probability of being a *BRCA* carrier[A] - Aged 40–69 years with a known *BRCA1* or *BRCA2* mutation *Offer as part of the population screening programme to women:* - Aged 70 years and over with a known *BRCA1* or *BRCA2* mutation *Consider annually for women:* - Aged 30–39 years at high risk of breast cancer but with a 30% or lower probability of being a *BRCA* carrier[A] - Aged 30–39 years who have not had genetic testing but have a greater than 30% probability of being a *BRCA* carrier[A] - Aged 30–39 years with a known *BRCA1* or *BRCA2* mutation
MRI surveillance	Do not offer at any age	*Offer annually to women:* - Aged 30–49 years who have not had genetic testing but have a greater than 30% probability of being a *BRCA* carrier[A] - Aged 30–49 years with a known *BRCA1* or *BRCA2* mutation
Risk-reducing mastectomy	Do not offer	- Should be raised as a risk-reducing strategy option with all women at high risk - Women considering this should have specialist genetic counselling
Risk-reducing oophorectomy	Do not offer	- Information should be provided as a potential risk-reducing strategy to women at high risk

The interventions here pertain to women with no personal history of cancer not yet eligible for inclusion in the NBSP and those with no mutation and at risk of or known *BRCA* mutation. Other recommendations are made in NICE CG164 for women with a personal history, of the age group eligible for the NHS breast screening programme and with other conditions/mutations conferring an increased risk of breast cancer. [A]NICE recommends that *BRCA* carrier probability be assessed using BOADICEA or MSS.

Despite there being no formal FH screening for all women with a FH of breast cancer, Evans *et al.* concluded it is possible that up to 5-6% of women age 46-49 years in a UK cohort may be eligible for additional screening according to the algorithm provided by NICE (213). It is not, however, clear if the same proportion would be eligible if the suggested alternative risk model, BOADICEA, were used to determine 10-year risk. Despite this, the proportion of women who would benefit from additional screening before the age of 50 has the potential to be significant. Referral for risk estimation depends on women seeking advice regarding their risk, or on general practitioners suggesting referral for patients whom they know have a FH. Of these women, identifying those who are indeed at increased risk is important. Any additional screening should have optimal sensitivity and specificity, to detect as many

cancers, as early as possible, whilst remaining cost effective. Accurate risk stratification not only allows patients to be offered beneficial screening, but allows them to make informed decisions about interventions such as risk-reducing surgery, should they be eligible. However, Evans et al. found the sensitivity of the NICE algorithm for moderate risk women to be 10.8% (95% CI, 3.5-23.4%), and for high risk 2.7% (0.3-11.5%). Specificity was 97.1% (96.6-97.6%) for moderate risk and 99.3% (99.1-99.5%) for high risk. The positive predictive value (PPV) for NICE moderate risk was 3.1%, and high risk 3.3% (213). Together, of the 37 cancers identified in the study, only 5 were accounted for in the moderate or high risk category. The poor sensitivity in this particular cohort (n=4,360) would suggest a need for re-evaluation of the NICE algorithm. Indeed, the study found that sensitivity increased to 54.1% if third degree relatives with breast cancer were included in FH, which NICE does not currently take into consideration (213). However, it is worth noting firstly, that the confidence intervals for sensitivity analysis in this project are wide, potentially due to the fact that just 14.8% of the cohort (649/4,360) had a family history of breast cancer in at least one SDR, and 9.4% (410/4,360) had a family history of at least one FDR. This reflects the fact that the cohort used in this study were selected from women invited for routine screening, rather than using a cohort enriched for women with a family history of breast cancer. It therefore perhaps doesn't provide the most relevant statistical estimates of the NICE guidelines. Nevertheless, the cost-effectiveness of additional screening for women deemed at sufficiently increased risk is important - avoiding over-screening of women should be balanced with achieving an acceptable level of sensitivity.

Of course, NICE CG164 is a guideline, and clinicians in specialist genetic centres will use their own clinical judgement, taking other factors into consideration when assessing a patient's individual risk of breast cancer. Nonetheless, the sensitivities calculated by Evans et al. are perhaps lower than would be expected for an assessment tool which is used nation-wide. For this reason, further research into the efficacy of the NICE algorithm in a separate cohort is useful, particularly as there was such a narrow age range used in the previously mentioned study.

1.5. Basis of this Research

It is clear from reviewing the literature that there is significant demand for a risk model for familial breast cancer which is both accurate, and simple to use in a clinical setting. Ultimately, this should inform clinical management of these patients, including screening. With any screening program, its sensitivity and specificity are important to ensure its cost-effectiveness. To our knowledge, when this study commenced, there had be no research attempting to validate the effectiveness of the NICE guidelines, used UK wide, in a cohort of women who have attended clinical genetics regarding their risk. In Tayside, there are the means to retrospectively analyse a large group of women who have attended clinical genetics and determine their NICE risk category and outcome over a period of time. This is a pilot study aiming to assess the effectiveness of NICE familial breast cancer guidelines in a Tayside cohort, and demonstrate methodology which could be expanded to other centres.

2. Aims and Objectives

2.1 Aims

The aims of this study are:

- To investigate how effective the NICE risk guidelines for familial breast cancer are at identifying women at an increased risk of breast cancer.
- To investigate the relative and absolute risk in a Tayside cohort of developing breast cancer depending on NICE risk category.
- To identify which elements of FH appear to be related to risk of invasive breast cancer.
- To investigate whether NICE guidelines identify women who will benefit from increased screening before entering into the NBSP i.e. those who develop breast cancer prior to age 50.

2.2. Objectives

To achieve the above aims, the following objectives will be carried out:

- A cohort of patient who have attended clinical genetics regarding risk of breast cancer will be identified and assigned a NICE risk category. Women who subsequently developed breast cancer will be identified.
- Based on the collected information, risk of breast cancer for women in each NICE risk category will be calculated.
- Statistical analysis will be performed to assess if any particular element of FH is significantly related to breast cancer risk.
- Age specific statistical analysis will be carried out to assess if women in higher NICE risk categories are in fact at significantly increased risk of breast cancer prior to age 50, at the beginning of the NBSP.

3. Methods

3.1. Approvals and data collection

3.1.1. Approvals

NHS Tayside Caldicott approval was granted to extract data pertaining to up to 1,500 women who attended the clinical genetics department from 2000-2010. This was to assess their risk of breast cancer according to NICE guidelines, and identify women who then developed breast cancer. In addition, permission was given to identify women within Tayside, but outwith the cohort, who developed breast cancer under age 50. Please see appendix 1 for a copy of this approval (*Caldicott reference number CSAppLL2349*). Ethics committee approval was not necessary for this study as there was no patient contact.

3.1.2. Inclusion/Exclusion Criteria

This was a retrospective, longitudinal cohort study consisting of patients referred for genetic counselling regarding FH of breast cancer from 2000-2010. Inclusion criteria included being a female age less than 50 at age of initial consultation, seen between 2000-2010, with no personal history of breast and/or ovarian cancer, and sufficient FH information available pertaining to the time of initial consultation. Patients were excluded if there was not sufficient recorded FH information or they did not attend their appointment.

3.1.3. Clinical Data Collection and Handling

Patients who met inclusion criteria were identified, and their FH information was collected from clinical genetics electronic and paper records at Ninewells Hospital and Medical School, NHS Tayside. The data was extracted from a combination of FH questionnaires, clinical pedigrees, clinical notes and official correspondence from or to the clinical genetics department. A list of clinical variables collected is provided in Appendix 2. Where possible, cancers reported within the family were confirmed by information in the clinical notes from the Information Services Division Scotland. After initial risk assessment according to NICE guidelines, women who were *BRCA*

mutation carriers were identified, again through electronic records in clinical genetics.

Women within the cohort who went on to develop breast cancer were identified through biopsy results from the Labcentre database within the pathology department at Ninewells hospital. The subtype of breast cancer was recorded, as well as date of biopsy diagnosis through the Integrated Clinical Environment data (ICE).

The Labcentre database was also used to identify all women who were diagnosed with breast cancer from 2000-2016, within NHS Tayside, who were under the age of 50 at the time of diagnosis. The CHI numbers for these patients were then used to interrogate the Ninewells clinical genetics database, to identify which of them had been seen in clinical genetics at any time. This information was used to establish the proportion of women developing breast cancer <50 years from 2000-2016 being detected by the FH screening programme.

All patient identifiable information was stored within the Ninewells Hospital computer network in NHS Tayside. To allow use of the data outwith this setting, it was pseudoanonymised by assignment of a study identification number, and stored within a password protected file, on a password protected laptop. Corresponding CHI and Study ID remained stored in the clinical genetics storage drive within the NHS Tayside network, separate from the pseudoanonymised information.

3.2. Assigning NICE risk category

All women were risk categorised *de novo* into low, moderate and high risk based on their FH information as outlined in the NICE guidelines on familial breast cancer (79). They were further assigned a sub-category based on the criteria presented in *Table 10* (section 1.4.7.). A woman with bilateral breast cancer counted as two individual breast cancer diagnoses, as suggested by NICE. Where a women had a FH of ovarian or male breast cancer and was therefore possibly eligible for more than one high risk category, category 13 was prioritised over category 12 which was prioritised over category 11. So, for example, a woman with a FH of bilateral breast cancer and male breast cancer would be categorised into category 13. Although clinically, NICE

guidance would be used alongside the clinician's judgement, for the purposes of research, the guidelines were interpreted in a literal sense.

3.3. Manchester Score Calculations

A Modified Manchester Score (MMS) was calculated for each patient using the ICR MMSS (210). This was chosen as a scoring system that is used by clinicians in Tayside and that doesn't require extensive information about familial cancers. This can be found in *Table 8* (section 1.4.6.). For example, a patient with one relative diagnosed with ovarian cancer <60 years old and two female relatives diagnosed with breast cancer between 50-59 years old would have a MMS of 13+4+4=21. The cut-off used clinically for suggested *BRCA* testing with the ICR MMSS is 17. The MMS also has the benefit of women who are related to a patient through a male relative being moved up one degree of relation – therefore unaffected intervening males are accounted for. Full information on how the MMS is applied can be found at www.icr.ac.uk (216).

3.4. *BRCA* mutation carriers

BRCA mutations present within the cohort were also identified as described in section 3.1.3. Where a *BRCA* mutation carrier was identified for whom their FH would have placed them in a moderate/low risk group, their risk was reassigned as high risk to best reflect how they would be managed clinically. In addition, women from a family with a known *BRCA* mutation who themselves tested negative were reassigned as low risk.

3.5. Invasive breast cancers and *in situ* carcinoma

Women who developed neoplastic breast disease were identified as described in section 3.1.3. There were women who were diagnosed with invasive breast cancer as well as *in situ* neoplastic disease. Limited analysis was performed on the *in situ* disease. DCIS and LCIS are widely regarded as cancer precursor lesions, however the natural course of these lesions is poorly understood (217). Although women identified with these lesions are usually offered surgical and systemic treatment, the risk factors determining progression to invasive breast cancer, and the benefit of treatment on mortality outcomes is uncertain (218). Only 20-50% of DCIS, which

accounts for the majority of *in situ* lesions will ever progress to invasive disease (219). If the aim of screening is to reduce mortality outcomes due to breast cancer, the significance of *in situ* disease is difficult to determine. Due to this uncertainty, very limited analysis was performed regarding *in situ* disease. The focus of this study was primarily on risk of invasive breast cancer.

3.6. Statistical Analysis

RR, OR, PPV, negative predictive value (NPV) and incidence were calculated using standard methods (220). The low risk group was used as the reference group. SPSS statistics software (221) was used for all other analyses of the data gathered in this study and was also used to generate charts and figures. For analyses generating a significance value (*p*-value), the conventional cut off of $p=<0.05$ was used to determine significance.

3.6.1. Percentage 10-year absolute risk calculation

A % 10-year absolute risk for women within the cohort at each NICE risk category, and also for *BRCA* mutation carriers was calculated. This was performed for risk between age 40-49 years and 50-59 years inclusive. The number of years follow up between those age ranges for all women in the cohort was summated (referred to as woman years of follow up). The % incidence of breast cancer per woman year of follow up, either between 40-49 or 50-59 was then calculated, and multiplied by 10 to give the approximate 10-year risk for that age range. % 5-year absolute risk would have been desirable as underlying population breast cancer risk will increase even in this short time period, however this was limited by cohort size and the number of cancers which occurred.

3.6.2. Chi-square test, Fischer's exact test, independent T-test and one-way analysis of variance

Categorical variables, such as development of cancer (yes or no) and affected mother (yes or no) were compared. Where there were sufficient numbers in the analysis, the Pearson Chi-square test was used, reporting asymptotic *p*-value (2-sided). Otherwise

a Fischer's exact test was used (*p*-value (2-sided)). The categorical variables analysed against development of cancer are shown in *Table 12.*

For the comparison of a continuous variable and a binomial categorical variable, and independent T-test was used. This was subject to Levene's test of variance to ensure normal distribution of data. Where there was a significant difference in Levene's test (i.e. p=<0.05), the reported *p*-value (2-sided) for the independent T-test does not assume equal variance of the continuous data. An example is the analysis of mean age of relative at diagnosis (continuous) and development of breast cancer (yes or no). The continuous variables analysed are shown in *Table 12.* These variables, along with the categorical variables were selected as broad descriptors of the cancer FH, representing similar family structures described the NICE guidance.

Where the means of a continuous variable were to be analysed for significance across a categorical variable with more than two groups (in this case, the three NICE risk categories), a one-way analysis of variance (ANOVA) was performed.

Table 12. Categorical and continuous variables selected for statistical analysis

Categorical variables analysed in Chi-square analysis	Continuous variables analysed in Independent T-test analysis
1 affected FDR or SDR ≥ 2 affected FDRs or SDRs ≥ 3 affected FDRs or SDRs ≥ 4 affected FDRs or SDRs Mother affected with breast cancer Sister affected with breast cancer Average age of relatives at diagnosis <40 Average age of relatives at diagnosis <50 Average age of relatives at diagnosis <60 Incidence of ovarian cancer in family Incidence of bilateral breast cancer in family Incidence of male breast cancer in family	Number of FDRs and SDRs affected in total Number of FDRs affected Average age of FDRs at diagnosis Number of SDRs affected Average age of SDRs at diagnosis Average age of all relatives at diagnosis

3.6.3. Kaplan-Meier Survival Analysis

Kaplan-Meier (KM) Survival Analysis was used to assess breast cancer development across different NICE risk categories and age ranges including total, <39 years, 40-49 years and 50-59 years of age. Time was measured in number of years follow up within that age range, and patients were censored at their completed time of follow up, or at breast cancer diagnosis. Separate analyses were performed to compare each NICE risk category. The high risk group were analysed both including and excluding *BRCA* carriers from the analysis as they are at known very high risk. KM survival curves were generated for selected sets of results.

3.7. Sample size calculation

It was important to address whether the study was adequately powered to assess the NICE guidelines ability to identify women at increased risk before age 50. To address this, a retrospective sample size calculation was performed. The sample size calculation was performed using methodology described by Jones, Carley and Harrison (222) for studies reporting categorical data i.e. diagnosis of invasive breast cancer. The following assumptions are made:

- The null hypothesis is that NICE guidelines do not effectively distinguish women at population, moderate (3-8%) and high (>8%) risk of breast cancer between the ages of 40-49.
- Type 1 error is to be avoided at the conventional level of 0.05 (pα).
- Type 2 error is to be avoided at the conventional level of 0.8 (pβ).
- The clinically important difference to be detected is the difference in % absolute risk from age 40-49 years, proposed by NICE, between the risk categories.

The following risk levels were used:

- Population risk was calculated using female invasive breast cancer rates per 100,000 reported for the year 2014 for the south east of Scotland. This data is available for download from http://www.isdscotland.org/ (223). The rate per 100,000 women per year for ages 40-44 and 45-49 was 114.4 and 201.6 respectively. This equates to % per woman per year risk of 0.114% for each year between age 40-44 and 0.2016% from age 45-49. Therefore, for one women the cumulative % risk from age 40-44 is 0.114 x 5 = 0.572%, and from 45-49 is 0.2016 x 5 = 1.008%. The % absolute risk between age 40-49 for a woman in the population of south east Scotland is therefore 0.572 + 1.008% = 1.580%. The population risk between ages 40-49 years used in the sample size calculation was therefore 1.580%.
- NICE states that a moderate risk woman has an absolute risk of 3-8% between ages 40-49. 3% was used as the risk for the sample size calculation so that the

sample size needed to detect the smallest possible difference could be calculated.

- NICE guidance suggests that high risk woman have an absolute risk of ≥8% between ages 40-49. Therefore, 8% was used as the risk for the sample size calculation. Again, this was to ensure that the sample size calculated was adequate to detect the smallest possible risk difference between the groups.

Initially the standardised difference between the proportions of expected breast cancers is calculated, as follows:

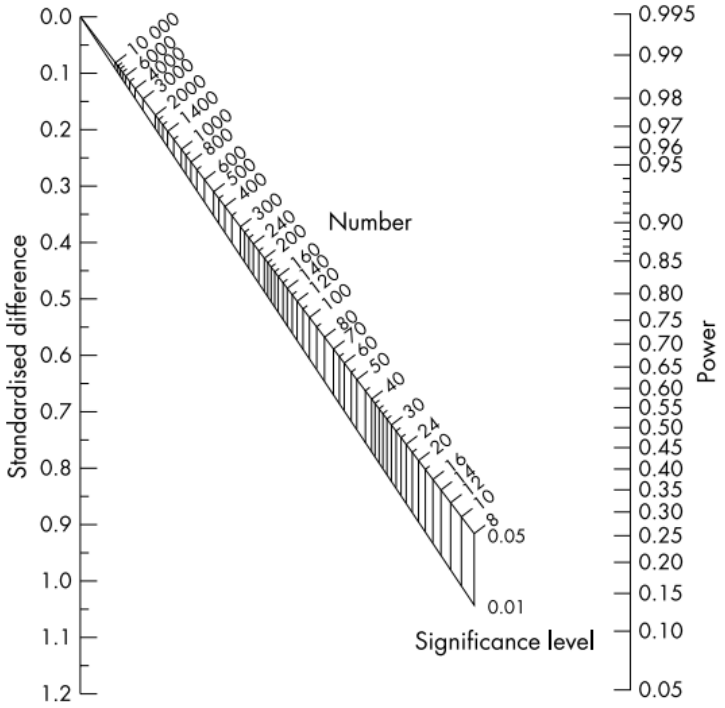
$$\text{Standardised difference} = p_1 - p_2 / \sqrt{(P(1-P))}$$

$$p_1 = \text{risk of breast cancer in higher risk group}$$
$$p_2 = \text{risk of breast cancer in lower risk group}$$

$$P = (p_1 + p_2)/2$$

The authors of the methodology then suggest that the standardized risk is used to work out the required sample size using a nomogram developed by Gore and Altman (224). This can be seen in *Figure 2*. A line is drawn from the calculated standard difference across to the designated pβ – in this case 0.8 – and the required sample size required at a pα level of 0.05 can be deduced.

Figure 2. Nomogram developed by Gore and Altman for sample size calculation (224) Image from an article by Jones, Carley and Harrison (222)



When the study commenced it was unknown what the number of available patients would be, therefore not possible to ascertain whether the study would be appropriately powered until completion. The calculation was therefore performed retrospectively to determine the sample size required to definitively validate the NICE guidelines ability to identify women at low/population, moderate or high risk of breast cancer. The result of this are as follows.

3.7.1. Sample size required to detect clinically significant difference between population and moderate risk group

$$\text{Standardised difference} = p_1\text{-}p_2/\sqrt{(P(1-P))}$$

p_1 = minimum risk of breast cancer between age 40-49 years at NICE moderate risk threshold = 3% or 0.03

p_2 = population risk of breast cancer between age 40-49 years = 1.580% or 0.0158

$$P = (p_1 + p_2)/2$$

$$P = \frac{(0.03 + 0.0158)}{2} = 0.0229$$

$$Standardised\ difference = \frac{0.03 - 0.0158}{\sqrt{0.0229(1 - 0.0229)}}$$

$$Standardised\ difference = \frac{0.0142}{\sqrt{0.0224}}$$

$$Standardised\ difference = \frac{0.0142}{0.14967}$$

$$Standardised\ difference = 0.095$$

Using the nomogram developed by Gore and Altman (224) (see *Figure 2,* section 3.7*)*, a standardized difference of 0.095 with a pβ of 0.8 and pα of 0.05 requires a population sample size of 4,000 subjects to give adequate power to the study. 4,000 patients with follow up between ages 40-49 would therefore be required to assess whether or not NICE guidelines identify women with at least a 3% risk of breast cancer between ages 40-49. There were 425 low, and 413 moderate risk (totalling 838) women with follow up between the ages of 40-49 in the study cohort, meaning this study is likely to be significantly underpowered to detect this clinical difference.

3.7.2. Sample size required to detect clinically significant difference between population and high risk group

$$\text{Standardised difference} = p_1\text{-}p_2/\sqrt{(P(1-P))}$$

p_1 = minimum risk of breast cancer between age 40-49 years at NICE high risk threshold = 8% or 0.08

p_2 = population risk of breast cancer between age 40-49 years = 1.580% or 0.0158

$$P = (p_{1+} p_2)/2$$

$$P = \frac{(0.08 + 0.0158)}{2} = 0.0958$$

$$Standardised\ difference = \frac{0.08 - 0.0158}{\sqrt{0.0958(1 - 0.0958)}}$$

$$Standardised\ difference = \frac{0.0642}{\sqrt{0.0866}}$$

$$Standardised\ difference = \frac{0.0642}{0.29428}$$

$$Standardised\ difference = 0.218$$

Using the nomogram developed, a standardized difference of 0.218 with a pβ of 0.8 and pα of 0.05, requires a population sample size of approximately 800 subjects to give adequate power to the study. 800 patients, with follow up between ages 40-49 would therefore be required to assess whether or not NICE guidelines identify women with at least an 8% risk of breast cancer, between ages 40-49. There are 425 low, and 294 high risk (totalling 719) women with follow up between the ages of 40-49 in this cohort, meaning this study is likely to be slightly underpowered to detect this clinical difference.

3.7.3. Sample size required to detect clinically significant difference between moderate and high risk group

$$\text{Standardised difference} = p_1\text{-}p_2/\sqrt{(P(1\text{-}P))}$$

p_1 = minimum risk of breast cancer between age 40-49 years at NICE moderate risk threshold = 3% or 0.03

p_2 = minimum risk of breast cancer between age 40-49 years at NICE high risk threshold = 8% or 0.08

$$P = (p_{1+}\ p_2)/2$$

$$P = \frac{(0.08 + 0.03)}{2} = 0.055$$

$$Standardised\ difference = \frac{0.08 - 0.03}{\sqrt{0.055(1 - 0.055)}}$$

$$Standardised\ difference = \frac{0.05}{\sqrt{0.0520}}$$

$$Standardised\ difference = \frac{0.05}{0.22804}$$

$$Standardised\ difference = 0.219$$

Using the nomogram, a standardized difference of 0.219 with a pβ of 0.8 and pα of 0.05, requires a population sample size of approximately 800 subjects to give adequate power to the study. 800 patients with follow up between ages 40-49 would therefore be required to assess whether or not NICE guidelines distinguish women at 3%, and women at 8% risk of breast cancer between ages 40-49. There are 413 moderate and 294 high risk (totalling 707) women with follow up between the ages of 40-49 in this cohort, meaning this study is likely to be slightly underpowered to detect this clinical difference.

4.0 Results

4.1. Descriptive Statistics

4.1.1. Cohort Characteristics

In total, 2,009 patient records were screened for eligibility in the study. 1,566 of these met the criteria of being a female under the age of 50, unaffected with breast cancer, and referred to clinical genetics regarding FH of breast cancer from 2000-2010. After this, 115 patient records were excluded as they did not attend their appointment. A further 38 were excluded as not enough information was present - FH information was regarded as suitable for use when acceptably detailed for at least 1st and 2nd degree relatives. Another 4 were excluded as they had been previously affected with ovarian cancer. Therefore, in total, 1,409 patients were eligible for inclusion in the study. The age at presentation to clinical genetics ranged from 16-49 years. The mean age was 36.75 (standard deviation (SD) ±7.855) years and the median was 38 years. All patients were referred to clinical genetics with a FH of breast cancer and/or ovarian cancer but without a personal history. The total number of years follow up time for all patients was 15,414 patient years. The mean number of years follow up was 10.935 (±3.3311) years with a median of 10.960 years.

4.1.2. Family history structures and cancer history

The variability in FH structure recorded from clinical notes is recorded in *Table 13.*

Table 13. Information regarding number of relatives within families in the cohort

	Mean	Range	
		Minimum	Maximum
Sisters	1.01	0	8
Daughters	0.54	0	5
Maternal aunts	1.40	0	9
Paternal aunts	1.04	0	10
Maternal half-sisters	0.03	0	5
Paternal half-sisters	0.02	0	5
Nieces	0.07	0	7

Table 14 gives an overview of the number and percentage of the cohort who had reported a history of breast cancer in various family members. The most common relative reported to have been affected was the mother (60.6%) followed by the maternal aunt (24.6%). 21.8% of the cohort had no history of breast cancer in a FDR, and 47.5% had no history of breast cancer in a SDR. Whilst 47.1% presented with only 1 affected relative, 48.8% presented with 2 or more relatives affected with breast cancer. 3.9% reported no FH of breast cancer, presenting instead with a FH of ovarian cancer. 10.8% of the cohort had a FH of ovarian cancer, and 1.8% had a history of breast cancer in a male relative. More detailed information regarding the frequency of responses can be found in *Table 15*.

Table 14. An overview of the reported cancer incidence in relatives of those within the cohort

	Incidence in relative amongst cohort (*N*=1409)	
	N	% of cohort
Breast cancer in Mother	854	60.6
Breast cancer in Sister	279	19.8
Breast cancer in Daughter	2	0.1
Breast cancer in Maternal aunt	346	24.6
Breast cancer in Paternal aunt	153	10.9
Breast cancer in Maternal grandmother	274	19.4
Breast cancer in Paternal grandmother	113	8.0
Breast cancer in Maternal half-sister	5	0.4
Breast cancer in Paternal half-sister	5	0.4
Breast cancer in Niece	1	0.1
Breast cancer in Male relative	26	1.8
Ovarian cancer in any relative	152	10.8
Bilateral breast cancer in 1st or 2nd degree relative	110	7.8

Table 15. Number of specific relatives diagnosed with cancer

	Frequency	% of cohort	Mean	Median	Range	
					Minimum	Maximum
Number of breast cancers among FDRs & SDRs						
0	55	3.9				
1	664	47.1				
2	471	33.4				
3	175	12.4	1.649	1.000	0	7
4	33	2.3				
5	7	0.5				
6	2	0.1				
7	2	0.1				
Number of breast cancers among FDRs						
0	307	21.8				
1	972	69.0	0.88	1.00	0	3
2	127	9.0				
3	3	0.2				
Mother affected with breast cancer						
Yes	790	56.1				
Bilateral	64	4.5	-	-	-	-
No	555	39.4				
Number of breast cancers among sisters						
0	1130	80.2				
1	252	17.9	0.22	0	0	3
2	24	1.7				
3	3	0.2				
Number of breast cancers among daughters						
0	1407	99.9	-	-	-	-
1	2	0.1				
Number of breast cancers among SDRs						
0	669	47.5				
1	484	34.4				
2	186	13.2				
3	59	4.2	0.77	1.00	0	7
4	7	0.5				
5	1	0.1				
6	2	0.1				
7	1	0.1				
Number of breast cancers among maternal aunts						
0	1062	75.4				
1	255	18.1	0.32	0.00	0.00	4
2	80	5.7				
3	11	0.8				
4	1	0.1				
Number of breast cancers among paternal aunts						
0	1256	89.1				
1	109	7.7				
2	31	2.2	0.15	0	0	7
3	9	0.6				
4	3	0.2				
7	1	0.1				
Maternal grandmother affected with breast cancer						
No	1135	80.6	-	-	-	-
Yes	258	18.3				
Bilateral	13	1.1				
Paternal grandmother affected with breast cancer						
No	1296	92.0	-	-	-	-
Yes	112	7.9				
Bilateral	1	0.1				
Number of family members affected with ovarian cancer						
0	1257	89.2				
1	124	8.8	0.13	0	0	4
2	23	1.6				
3	2	0.1				
4	3	0.2				
Total number of relatives affected with bilateral breast cancer						
0	1299	92.2	0.08	0	0	2
1	107	7.6				
2	3	0.2				

Not shown in *Table 15* is the number of women who reported a sister or aunt with bilateral breast cancer. 7.8% of the cohort reported bilateral breast cancer in at least one 1st or 2nd degree relative. There were 11 cases of bilateral breast cancer in a sister. 28 women reported at least 1 maternal aunt with bilateral breast cancer, and 14 women reported a paternal aunt with bilateral breast cancer. There were no reports of bilateral breast cancers occurring in the daughter, half-sister or niece of any of the women in the study. Of the 26 male breast cancers reported, 7 occurred in a father, 1 in a brother, 6 in a maternal uncle, 3 in a paternal uncle, 2 in a maternal grandfather and 2 in a paternal grandfather. 5 of the male breast cancers occurred in a more distant relative.

Information regarding the ages of diagnosis of the female breast and ovarian cancers, for the whole cohort, is detailed in *Table 16*. The age at bilateral breast cancer diagnosis reflects the mean age at diagnosis of the two cancers.

Table 16. Average ages of cancer diagnosis within family members of the entire cohort

	Mean (±SD)	Median	Range	
			Minimum	Maximum
Age of total FDRs and SDRs at breast cancer diagnosis	50.308 (±10.817)	50	19.50	97.00
Age of ovarian cancer diagnosis	55.376 (±13.740)	53.75	22.00	87.00
Age of bilateral breast cancer diagnosis	52.864 (±10.596)	51.750	32.0	79.0
Age of FDR(s) at breast cancer diagnosis	48.810 (±10.991)	48.00	21.00	87.00
Age mother at breast cancer diagnosis	51.321 (±11.098)	50.75	21.00	87.00
Age sister(s) at breast cancer diagnosis	41.6261 (±7.324)	42.00	23.00	63.00
Age daughter(s) at breast cancer diagnosis	27.500 (±0.707)	27.500	27.00	28.00
Age of SDR(s) at breast cancer diagnosis	55.509 (±12.640)	55.00	23.00	97.00
Age of maternal aunt(s) at breast cancer diagnosis	52.5614 (±11.230)	51.00	25.00	82.00
Age of paternal aunt(s) at breast cancer diagnosis	56.029 (±10.740)	55.00	32.00	89.00
Age of maternal grandmother at breast cancer diagnosis	58.620 (±14.539)	58.00	29.00	92.00
Age of paternal grandmother at breast cancer diagnosis	59.2611 (±14.219)	59.00	35.00	97.00

4.1.3. NICE risk categories

After assigning NICE risk categories based on FH, 506 (35.9%) were low risk, 523 (37.1%) were moderate risk, and 380 (27.0%) were high risk according to NICE guidance. There were adjustments made to this made based on the results of *BRCA* testing, which were collected following initial NICE risk assignment (see section 4.1.5. for details). The final numbers in each group were therefore 505 (35.8%) low risk, 522 (37%) moderate risk and 382 (27.1%) high risk (including 12 *BRCA1* and 10 *BRCA2* carriers). The corresponding mean age at presentation in these groups was 37.32 (SD±7.741), 36.60 (±7.868) and 36.19 (±7.957). One-way ANOVA demonstrated no significant difference at age of presentation between the three groups (p=0.092). The breakdown of the FH subtypes which contribute to each risk category, before adjustments for *BRCA* mutation status, can be found in *Table 17* (those in the low risk group did not meet any of the criteria shown). Of those in the moderate risk category, the majority (58.3%) had a FH of two FDRs or SDRs diagnosed with breast cancer, at an average age of older than 50. The most common reason for being assigned a high risk was having two FDRs or SDRs diagnosed with breast cancer, at an average age of younger than 50, with at least one affected FDR (32.1% of all high risk patients). It is worth noting that not all patients with a FH of ovarian cancer or male breast cancer are automatically assigned a high risk category if they do not meet the additional criteria outlined.

Table 17. NICE risk categories and specific criteria assigned to the cohort (before *BRCA* mutation results)

	Frequency	% of cohort
Moderate Risk		
One FDR diagnosed with breast cancer at younger than age 40 years	172	12.2
Two first-degree or SDRs diagnosed with breast cancer at an average age of older than 50 years	305	21.6
Three first-degree or SDRs diagnosed with breast cancer at an average age of older than 60 years	46	3.3
	523	37.1
High Risk		
At least the following female breast cancers only in the family:		
Two first-degree or SDRs diagnosed with breast cancer at younger than an average age of 50 years (at least one must be a FDR)	122	8.7
Three first-degree or SDRs diagnosed with breast cancer at younger than an average age of 60 years (at least one must be a FDR)	46	3.3
Four relatives diagnosed with breast cancer at any age (at least one must be a FDR)	14	1.0
Families containing one relative with ovarian cancer at any age and, on the same side of the family:		
One FDR (including the relative with ovarian cancer) or SDR diagnosed with breast cancer at younger than age 50 years.	41	2.9
Two first-degree or SDRs diagnosed with breast cancer at younger than an average age of 60 years.		
Another ovarian cancer at any age.	33	2.3
Families affected by bilateral cancer (each breast cancer has the same count value as one relative):	24	1.7
One FDR with cancer diagnosed in both breasts at younger than an average age 50 years.		
One first-degree or SDR diagnosed with bilateral cancer and one first or SDR diagnosed with breast cancer at younger than an average age 60 years.	28	2.0
	48	3.4
Families containing male breast cancer at any age and, on the same side of the family, at least:		
One first-degree or SDR diagnosed with breast cancer at younger than age 50 years.		
Two first-degree or SDRs diagnosed with breast cancer at younger than an average age of 60 years.	6	0.4
	11	0.8
	380	27.0

4.1.4. Modified Manchester Scores

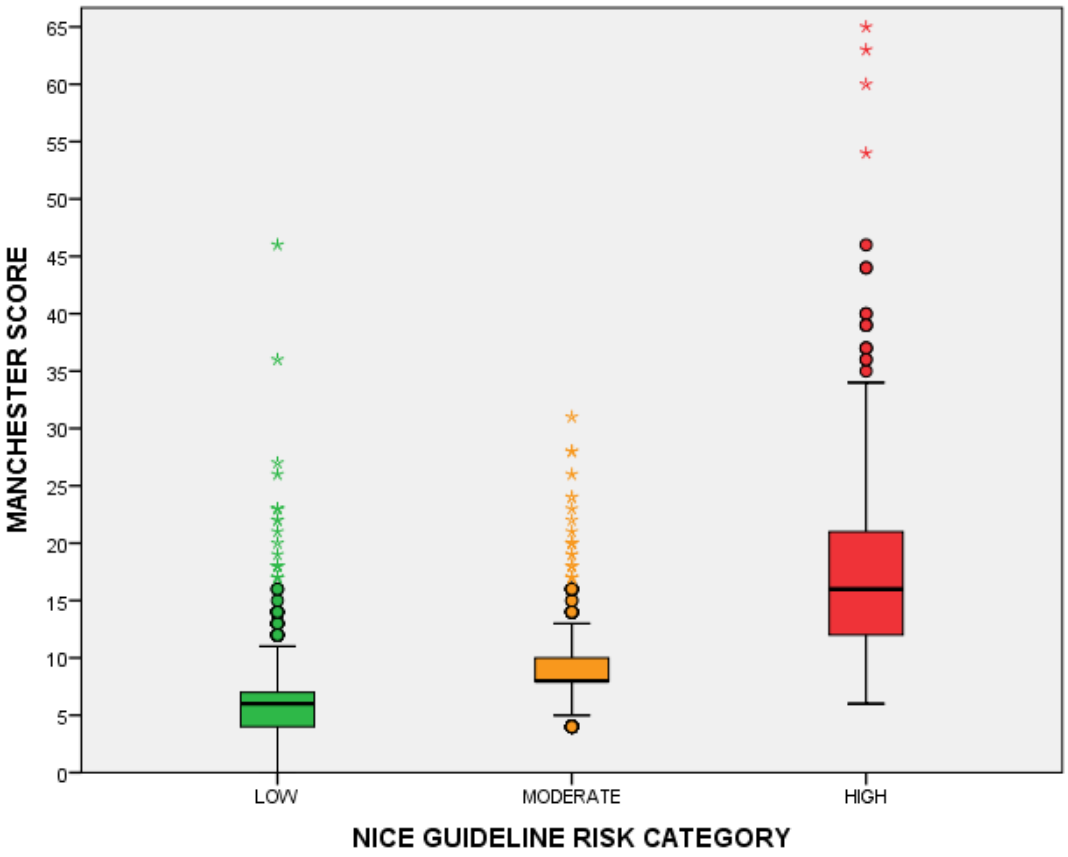
The mean MMS for the entire cohort was 10.47 (SD±7.402) with a median of 8.00. Scores ranged from 0-65. 95.5% of the cohorts' MMS fell within the range of 0-25. Scores meeting the threshold for *BRCA* testing i.e. ≥17, accounted for 15.6% of the cohort, equal to 220 patients.

For the low risk group, the mean MMS was 6.57 (±5.01). The median was 6.00 and scores ranged from 0-46. 95.1% of the scores in the low risk group ranged from 0-16, and 99.2% ranged from 0-23. Scores ≥17, accounted for 4.9% of the low risk group i.e. 25 patients. The scores of 17 and above, including the exceptionally high scores, can be accounted for due to 1) the fact that the MMSS moves affected female relatives linked through an unaffected male up one degree of relation, something not stated in the NICE risk evaluation 2) the relatively high value the MMS assigns to ovarian cancer and male breast cancer in patients for which the FH does not otherwise meet NICE risk criteria.

In the NICE moderate risk group, the mean MMS was 9.03 (±3.76) and the median score was 8.00. The scores in this group ranged from 4-31. 95.4% of the MMS ranged from 4-16. Scores ≥17 therefore accounted for 24 patients (4.6%) in the moderate risk group, equal to 1.7% of the entire cohort.

For the NICE high risk group, the mean MMS was 17.61 (±8.68), with a median score of 16.0. Scores ranged from 6-65, with 93.5% of the scores ≥10. Scores ≥17, i.e. the threshold for offering *BRCA* testing according to the MMSS, made up 44.6% of the high risk group. This equates to 169 high risk patients, i.e. 12.0% of the entire cohort. The boxplot of these scores can be seen in *Figure 3*, which also demonstrates the significant outliers.

Figure 3. Boxplot of MMS by NICE risk category, with significant outliers shown.



4.1.5. *BRCA* mutation testing and results

62 patients in the cohort had been tested for a *BRCA* mutation. 34 of these patients had a family member who had tested positive previously. 33 had a MMS ≥17. One of the patients who was tested had neither a score ≥17 nor a family member who was *BRCA* positive. Of those tested, 22 patients in total were positive for a mutation – 12 patients had a mutation in *BRCA1* and 10 patients had a mutation in *BRCA2*. The mean MMS for *BRCA* carriers was 23.50 (SD±12.87). 15 of the 22 had an MMS of 17 or above.

As a result of acquiring data on *BRCA* testing, some patients changed risk category, as those with a *BRCA* mutation are automatically high risk. Some patients' risk category was also lowered if they tested negative for a known familial *BRCA* mutation, in order to reflect how the patient would be regarded clinically. The final NICE risk category assignments, using both FH, and *BRCA* test results, can be seen in *Table 18*. The finalised figures, incorporating *BRCA* test result, were used for further analysis.

Table 18. Original and final risk categorisation

	FH alone		FH and *BRCA* test result	
	N	% of cohort	*N*	% of cohort
Low risk	506	35.9	505	35.8
Moderate risk	523	37.1	522	37.0
High risk	380	27.0	382	27.1
Total	1409	100	1409	100

Modified Manchester Score and BRCA mutations

Independent T-test analysis was not appropriate, as the MMS did not meet the assumption of being normally distributed that is required for this test to be appropriate. Fischer's exact test was performed to determine whether the cut off of a MMS of 17 for *BRCA* testing suggested by Institute of Cancer Research predicted mutation in the cohort. The results of this analysis can be seen in *Table 19.* For having any *BRCA* mutation and for having either a *BRCA1* or *BRCA2* mutation, a MMS of ≥17 was significantly predictive. In addition, the sensitivity and specificity, as well as the AUROC for a MMS ≥17 and mutation status is demonstrated.

Table 19. Fischer's exact test, sensitivity, specificity and AUROC for MMS ≥17 and carrying a *BRCA* mutation

	MMS ≥17 *(N)*		
	Yes	No	*P*-value (Fischer's exact test)
BRCA1 mutation carrier	9	3	<0.001
Non-mutation carrier	211	1186	
BRCA2 mutation carrier	6	4	0.002
Non-mutation carrier	214	1185	
BRCA1 or *BRCA2* mutation carrier	15	7	<0.001
Non-mutation carrier	205	1182	

	Sensitivity (%) (95% CI)	Specificity (%) (95% CI)	AUROC (95% CI)	AUROC *p*-value
BRCA1	75.00 (42.84-93.31)	84.90 (82.89-86.71)	0.799 (0.657-0.942)	<0.001
BRCA2	60.00 (27.37-86.31)	84.71 (82.68-86.53)	0.724 (0.542-0.905)	0.015
BRCA1 or BRCA2	68.18 (45.12-85.27)	85.22 (83.22-87.02)	0.767 (0.652-0.882)	<0.001

4.1.6. Number of years follow up by age range and risk category

The number of women years of follow up by age bracket and NICE risk category is shown in *Table 20* and represents the figures used for later incidence and risk calculations. *BRCA* carriers are shown separately.

Table 20. Number of women years of follow up for each age range and NICE risk category.

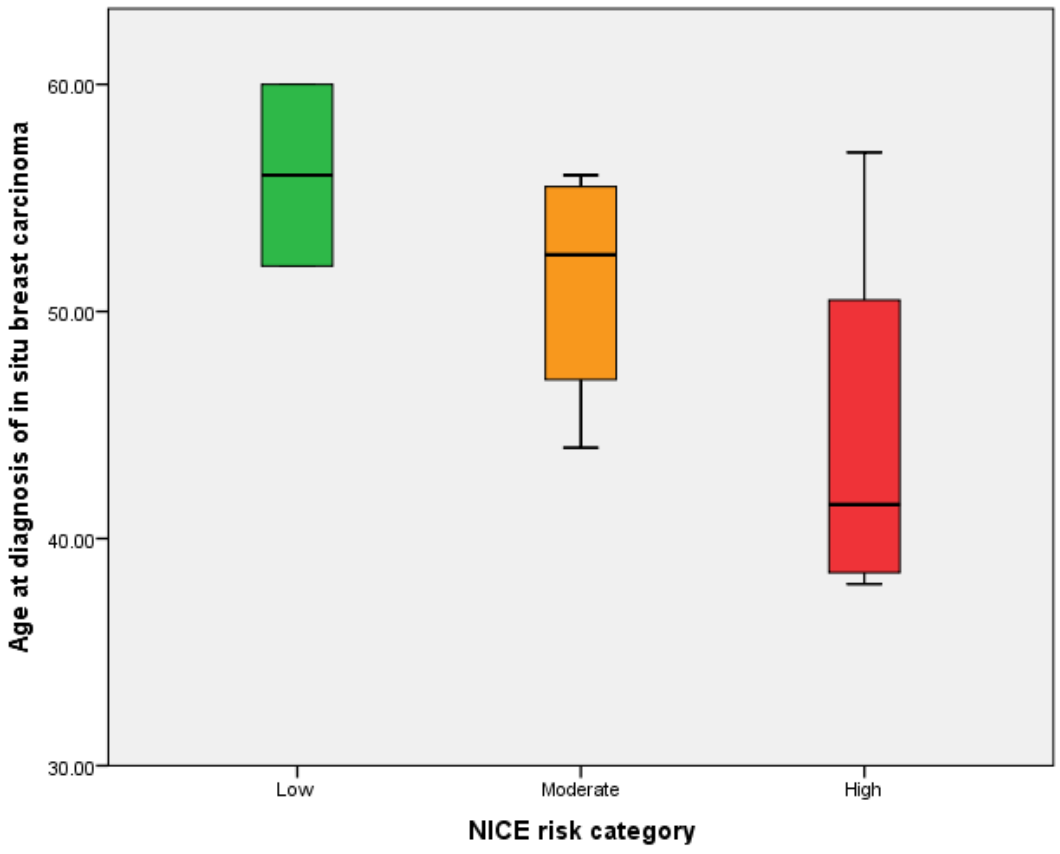
		<39 years	**40-49 years**	**50-59 years**	**>60 years**	**Total**
Low risk		1,871	2,430	1,239	91	5,631
Moderate risk		2,131	2,384	1,135	69	5,737
High risk	Non-*BRCA* carriers	1,414	1,609	757	52	3,813
	BRCA carriers	139	75	19	0	233
Total		5,555	6,498	3,150	212	**15,414**

4.1.7. Cancer development in the cohort

In Situ Carcinoma Development

10 women in the cohort developed *in situ* carcinoma. There were 9 cases of DCIS and 1 case of LCIS. None of the women were *BRCA* mutation carriers. Age of diagnosis ranged from 38-60 years. The mean age of diagnosis was 49.5 years (SD±7.807), and median age was 51 years. 4 occurred in women in the high risk group, 4 in the moderate risk group and 2 in the low risk group. There was no significant difference in the likelihood of developing *in situ* disease between the low and moderate (Fischer's exact p=0.687), low and high (p=0.411), or moderate and high (p=0.728) risk groups. A boxplot of the age at diagnosis can be seen in *Figure 4.* The difference in age of diagnosis between NICE risk groups was not significant on one-way ANOVA (p=0.211). However, the numbers in each group are small and the analysis is likely to be significantly underpowered.

Figure 4. Ages of in situ carcinoma diagnosis by NICE risk group



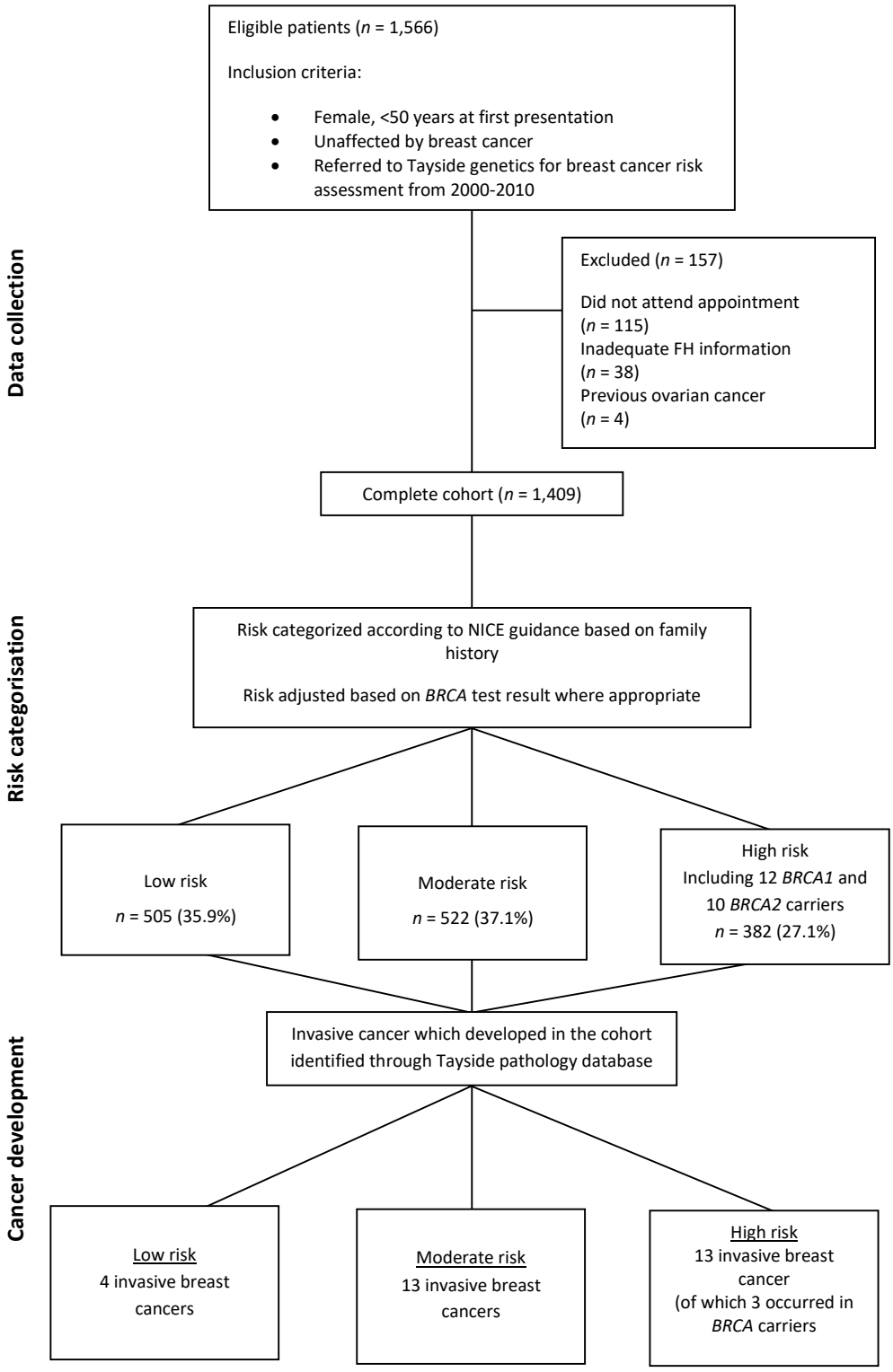
Invasive Cancer Development

Of the 1,409 women in the cohort, 30 developed an invasive cancer in the period between attending clinical genetics and May 2016. These patients first presented to clinic genetics at a mean age of 40.5 years (SD±5.76), with a median age of 41, ranging from age 29-49. The mean age of breast cancer development in the cohort was age 48.9 (±6.72), median 49.5, and ranged from age 32-59 years. 3 (10%) of the cancers occurred in a *BRCA* mutation carrier – these occurred at ages 40, 41 and 53. The mean time from first contact with clinical genetics to the development of invasive cancer was 8.01 years (±3.39), and ranged from 1.25-14.97 years. In total, the women who developed invasive breast cancer had cumulative number of years follow up of 360.3 years, mean 12.01 (±2.91), and range 6.6-16.40.

4.1.8. Cohort summary

Figure 5 shows a consort diagram summarising the data collection process and basic descriptive characteristics of the cohort.

Figure 5. Consort diagram overviewing data collection and final cohort



4.2. Risk Analysis

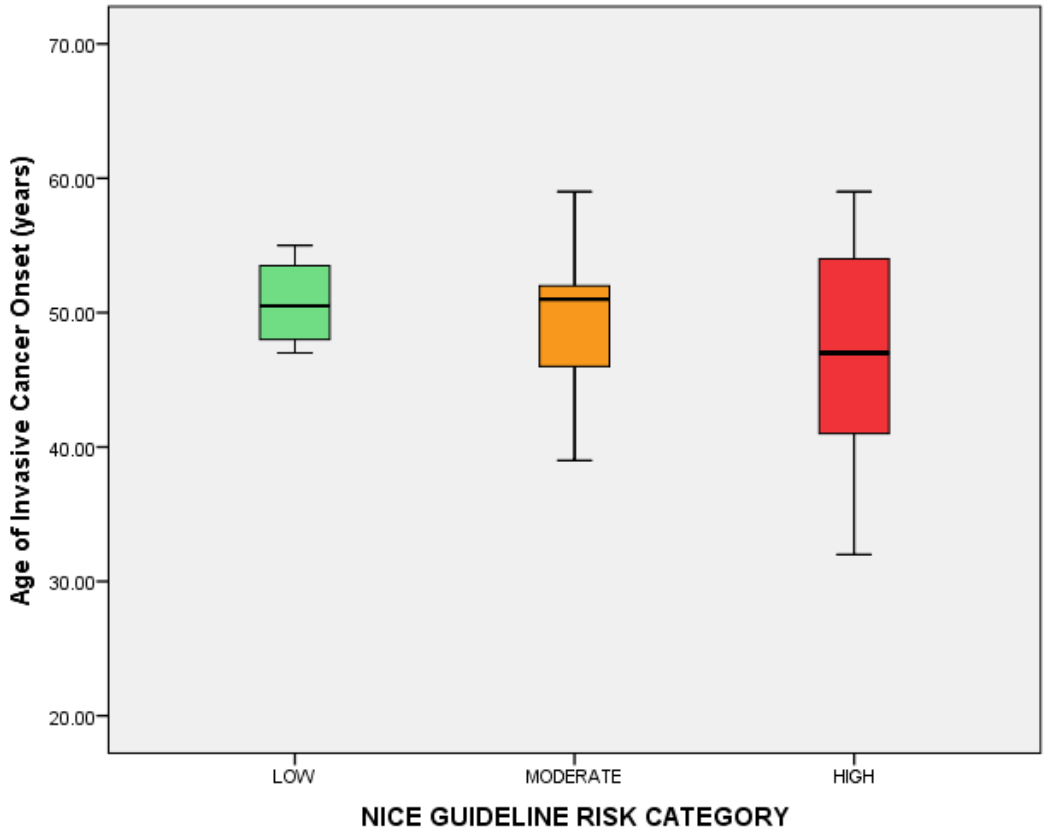
4.2.1. Mean age of cancer development by NICE risk category

Of the 30 cancers which developed, 4 occurred in the low risk group, 13 occurred in the moderate risk group, and a further 13 occurred in the high risk group. Mean age of cancer diagnosis for each group is described in *Table 21*, and is plotted in *Figure 6*. Age of cancer diagnosis appears to decrease with increasing risk category. This however did not reach statistical significance on one-way ANOVA (p=0.578).

Table 21. Age of invasive cancer diagnosis statistics by NICE category (years)

	Mean (±SD)	Median	Range
Low Risk	50.8 (±3.5)	50.5	47-55
Moderate Risk	49.8 (±5.6)	51.0	39-59
High Risk	47.5 (±8.4)	47.0	32-59

Figure 6. Boxplot of age of invasive cancer diagnosis by NICE risk category



4.2.2. Independent T-test analysis

Independent T-test analysis was performed to compare continuous variables with invasive breast cancer development, overall and across different age groups, both including and excluding *BRCA* mutation carriers. The full results of the analysis can be found in *Table 22.* It should be noted that the significance values presented are uncorrected for multiple testing.

4.2.3. Pearson Chi-square/Fischer's exact test

Fischer's exact test was performed to compare the collected categorical variables with invasive breast cancer development overall and across different age groups, both including and excluding *BRCA* mutation carriers. The full results of the analysis can be found in *Table 23.* Again, it should be noted that the significance values presented are uncorrected for multiple testing.

Table 22. Results of independent t-test analysis of continuous variables with breast cancer development

	Variable	Excluding *BRCA* carriers	Including *BRCA* carriers
		P-value (2-tailed)	*P*-value (2-tailed)
Overall cancer diagnoses (*N*=30)	Number of FDRs and SDRs affected in total	**0.004**	**0.003**
	Number of FDRs affected	0.60	**0.024**
	Average age of FDRs at diagnosis	0.088	0.269
	Number of SDRs affected	0.240	0.112
	Average age of SDRs at diagnosis	0.644	0.618
	Average age of all relatives at diagnosis	0.088	0.436
Cancer diagnoses age ≤39 years (*N*=3)	Number of FDRs and SDRs affected in total	0.189	0.192
	Number of FDRs affected	0.498	0.508
	Average age of FDRs at diagnosis	0.343	0.347
	Number of SDRs affected	0.086	0.090
	Average age of SDRs at diagnosis	0.337	0.326
	Average age of all relatives at diagnosis	0.565	0.557
Cancer diagnoses age 40-49 years (*N*=12)	Number of FDRs and SDRs affected in total	0.111	0.096
	Number of FDRs affected	0.203	0.436
	Average age of FDRs at diagnosis	**0.006**	0.089
	Number of SDRs affected	0.144	0.057
	Average age of SDRs at diagnosis	0.729	0.384
	Average age of all relatives at diagnosis	0.097	0.565
Cancer diagnoses age 50-59 years (*N*=15)	Number of FDRs and SDRs affected in total	**0.038**	**0.036**
	Number of FDRs affected	0.086	0.077
	Average age of FDRs at diagnosis	0.649	0.681
	Number of SDRs affected	0.779	0.668
	Average age of SDRs at diagnosis	0.807	0.675
	Average age of all relatives at diagnosis	0.706	0.744

Table 23. Results of Pearson Chi-square/Fischer's exact test of categorical variables with breast cancer development

	Variable	Excluding *BRCA* carriers	Including *BRCA* carriers
		P-value (2-tailed)	*P*-value (2-tailed)
Overall cancer diagnoses (*N*=30)	1 affected FDR or SDR	1.000	1.000
	≥ 2 affected FDRs or SDRs	**0.003**	**0.003**
	≥ 3 affected FDRs or SDRs	0.053	**0.039**
	≥ 4 affected FDRs or SDRs	0.205	0.228
	Mother affected with breast cancer	1.000	0.851
	Sister affected with breast cancer	0.465	0.643
	Average age of relatives at diagnosis <40	0.856	0.996
	Average age of relatives at diagnosis <50	0.083	0.275
	Average age of relatives at diagnosis <60	0.817	0.656
	Incidence of ovarian cancer in family	0.514	0.764
	Incidence of bilateral breast cancer in family	0.459	0.504
	Incidence of male breast cancer in family	0.077	0.104
Cancer diagnoses age ≤39 years (*N*=3)	1 affected FDR or SDR	-[a]	-[a]
	≥ 2 affected FDRs or SDRs	0.116	0.117
	≥ 3 affected FDRs or SDRs	0.392	0.398
	≥ 4 affected FDRs or SDRs	1.000	1.000
	Mother affected with breast cancer	0.565	0.565
	Sister affected with breast cancer	0.488	0.484
	Average age of relatives at diagnosis <40	0.889	0.886
	Average age of relatives at diagnosis <50	0.253	0.151
	Average age of relatives at diagnosis <60	1.000	0.487
	Incidence of ovarian cancer in family	1.000	0.710
	Incidence of bilateral breast cancer in family	0.214	0.217
	Incidence of male breast cancer in family	0.051	0.054
Cancer diagnoses age 40-49 years (*N*=12)	1 affected FDR or SDR	1.000	1.000
	≥ 2 affected FDRs or SDRs	0.059	0.084
	≥ 3 affected FDRs or SDRs	0.186	0.101
	≥ 4 affected FDRs or SDRs	1.000	1.000
	Mother affected with breast cancer	0.099	0.141
	Sister affected with breast cancer	0.226	0.139
	Average age of relatives at diagnosis <40	0.547	0.999
	Average age of relatives at diagnosis <50	0.117	0.397
	Average age of relatives at diagnosis <60	1.000	1.000
	Incidence of ovarian cancer in family	1.000	0.378
	Incidence of bilateral breast cancer in family	0.175	0.237
	Incidence of male breast cancer in family	1.000	1.000
Cancer diagnoses age 50-59 years (*N*=15)	1 affected FDR or SDR	1.000	1.000
	≥ 2 affected FDRs or SDRs	0.108	0.069
	≥ 3 affected FDRs or SDRs	0.248	0.268
	≥ 4 affected FDRs or SDRs	0.070	0.079
	Mother affected with breast cancer	0.422	0.602
	Sister affected with breast cancer	0.044	0.094
	Average age of relatives at diagnosis <40	0.981	0.994
	Average age of relatives at diagnosis <50	0.796	1.000
	Average age of relatives at diagnosis <60	1.000	1.000
	Incidence of ovarian cancer in family	0.385	0.393
	Incidence of bilateral breast cancer in family	0.617	0.622
	Incidence of male breast cancer in family	0.211	0.239

[a]Uncalculatable due to too few numbers

4.2.4. Modified Manchester Score

There is a trend towards higher MMS for women who developed invasive breast cancer compared to women who did not, both overall and at difference age ranges. However, area under the receiver operating curve (ROC) analysis assessing MMS with a cut-point of 17 and risk of developing breast cancer did not reach significance. This suggests no predictive ability of MMS for breast cancer at any age range. The mean MMS for each risk category and *p*-values for the AUROC analysis can be seen in *Table 24.*

Table 24. Mean MMS and AUROC for invasive breast cancer overall and by age range.

	Mean MMS (±SD)			
	Invasive breast cancer			
Excluding *BRCA* mutation carriers	Yes	No	AUROC	*P*-value
Overall	11.37 (±5.712)	10.25 (±7.121)	0.519	0.734
Age <39 years	14.67 (±5.033)	10.26 (±7.098)	0.593	0.578
Age 40-49 years	10.70 (±5.293)	10.26 (±7.111)	0.526	0.773
Age 50-59 years	11.14 (±6.249)	10.25 (±7.121)	0.498	0.978
Including *BRCA* mutation carriers				
Overall	11.77 (±5.746)	10.45 (±7.429)	0.522	0.674
Age <39 years	14.67 (±5.033)	10.47 (±7.401)	0.589	0.595
Age 40-49 years	11.75 (±5.786)	10.46 (±7.414)	0.548	0.570
Age 50-59 years	11.20 (±6.026)	10.45 (±7.429)	0.489	0.884

4.2.5. Frequency and percentage 10-year absolute risk of breast cancer

The frequency of cancer diagnosis in the cohort by NICE risk category and age range can be seen in *Table 25*. *BRCA* mutation carriers are shown both separately and included in the high risk group. The % 10-year absolute risk, based on women years of follow up in the cohort are shown for age ranges 40-49 and 50-59 years. Across all age ranges, the fewest cancer diagnoses occurred in the low risk group and this corresponded with the smallest % 10-year absolute risk. The highest absolute risk between the ages of 40-49 was in the high risk group, both including and excluding *BRCA* carriers, though *BRCA* carriers had the highest absolute risk overall in this age range. In the 50-59 years group, the moderate risk group had the highest % absolute risk, at 7.05%.

Table 25. Cancer diagnosis by age range and NICE category. % 10-year risk calculated based on number of women years of follow up, shown in *Table 21*.

	***BRCA* carriers separate**					***BRCA* carriers included**				
	N	Number of invasive cancers (% 10-year absolute risk (95% CI))				*N*	Number of invasive cancers (% 10-year absolute risk (95%CI))			
		Age range (years)					Age range (years)			
		Overall	**<39**	**40-49**	**50-59**		**Overall**	**<39**	**40-49**	**50-59**
Low Risk	505	4	0	2 (0.82% (0.72-0.94))	2 (1.61% (1.42-1.83))	505	4	0	2 (0.82% (0.72-0.94))	2 (1.61% (1.42-1.83))
Moderate Risk	522	13	1	4 (1.68% (1.53-1.83))	8 (7.05% (6.78-7.31))	522	13	1	4 (1.68% (1.53-1.83))	8 (7.05% (6.78-7.31))
High Risk	360	10	2	4 (2.49% (2.28-2.70))	4 (5.28% (4.93-5.64))	382	13	2	6 (3.56% (3.34-3.80))	5 (6.44% (6.10-6.78))
***BRCA* carriers**	22	3	0	2 (26.67% (17.98-37.63))	1 (52.63% (31.71-72.67))					
Total	1409	30	3	12	15	1409	30	3	12	15

4.2.6. Relative risks and odds ratios for NICE risk categories

The RR and OR for the moderate and high risk groups, with comparison to the low risk group are shown in *Table 26*. *BRCA* carriers are shown both as separate and included in the high risk group. Also shown is the RR and OR associated with meeting any of the NICE criteria, i.e. moderate or high risk. Other than being a *BRCA* carrier, the highest RR and OR for developing invasive cancer are in the high risk group, both including and excluding *BRCA* mutation carriers.

Table 26. RR and OR for NICE risk categories

	Relative risk (95% CI)	**Odds ratio (95% CI)**
Moderate Risk	3.14 (1.03-9.58)	3.20 (1.04-9.88)
High Risk (including *BRCA* carriers)	4.30 (1.41-13.07)	4.410 (1.43-13.64)
High Risk (excluding *BRCA* carriers)	3.51 (1.11-11.10)	3.58 (1.11-11.50)
***BRCA* carriers**	17.22 (1.10-72.29)	19.78 (4.13-94.63)
Moderate or high risk	3.63 (1.27-10.35)	3.71 (1.29-10.69)

4.2.7. Sensitivity, specificity, positive predictive value and negative predictive value of NICE risk categories

The sensitivity, specificity, PPV and NPV of the risk categories are shown in *Table 27*. These were calculated for moderate and high risk separately as compared to low risk, and in addition with the moderate and high risk categories grouped together. The sensitivity of being in any increased risk group was 86.67% however the specificity was much poorer at 36.33%.

Table 27. Sensitivity, specificity, PPV and NPV for NICE risk categories as compared with the low risk group.

	Sensitivity (%) (95% CI)	Specificity (%) (95% CI)	PPV (%) (95% CI)	NPV (%) (95% CI)
Moderate	76.47 (49.8-92.17)	49.6 (46.5-52.7)	2.49 (1.39-4.33)	99.2 (97.8-99.7)
High	76.47 (49.76-92.17)	57.59 (54.22-68.90)	3.40 (1.90-5.90)	99.21 (97.84-99.75)
Moderate & High	86.67 (68.36-95.64)	36.33 (33.80-38.94)	2.88 (1.93-4.25)	99.21 (97.84-99.75)

4.2.8. Area under the receiver operating curve for NICE risk categories

AUROC was analysed to determine the accuracy of NICE risk categories at determining the likelihood of developing breast cancer, at different age ranges, for different NICE risk categories, as compared to women in the NICE low risk category. The results can be seen in *Table 28*. Being in the moderate or high risk group combined, across total follow up time, predicted significantly better than chance the risk of developing breast cancer (p=0.031), when *BRCA* carriers are included in the analysis. This significance was lost on exclusion of *BRCA* mutation carriers. Being in the high risk group, including *BRCA* mutation carriers, also predicted significantly better than chance the risk of developing breast cancer across total follow up time (p=0.016). This significance was again, lost on exclusion of *BRCA* mutation carriers. The receiver operating curves (ROC) for the two significant results are shown in *Figure 7* and *Figure 8*.

Table 28. AUROC for varying age ranges and NICE risk categories.

	BRCA carriers excluded		*BRCA* carriers included	
	AUROC (95% CI)	*P*-value	AUROC (95% CI)	*P*-value
Moderate or high risk				
Overall	0.609 (0.514-704)	0.052	0.615 (0.526-0.704)	**0.031**
<39 years	0.681 (0.472-890)	0.277	0.680 (0.469-0.890)	0.282
40-49 years	0.582 (0.418-0.746)	0.371	0.597 (0.452-0.742)	0.246
50-59 years	0.612 (0.481-0.742)	0.150	0.615 (0.490-0.740)	0.125
Moderate risk				
Overall	0.630 (0.506-0.753)	0.066	0.630 (0.507-0.754)	0.065
<39 years	0.746 (0.457-1.000)	0.395	0.746 (0.458-1.000)	0.394
40-49 years	0.579 (0.357-0.802)	0.502	0.580 (0.358-0.802)	0.499
50-59 years	0.648 (0.493-0.802)	0.108	0.648 (0.494-0.802)	0.107
High risk				
Overall	0.649 (0.510-0.789)	0.055	0.670 (0.549-0.791)	**0.016**
<39 years	0.791 (0.621-0.960)	0.155	0.785 (0.612-0.959)	0.163
40-49 years	0.625 (0.405-0.844)	0.292	0.662 (0.484-0.840)	0.115
50-59 years	0.626 (0.406-0.845)	0.289	0.645 (0.449-0.842)	0.186

Figure 7. ROC for invasive cancer diagnosis over total patient follow up time, comparing the combined NICE moderate and high risk group (including *BRCA* carriers) to the NICE low risk group (AUROC=0.615 (0.526-0.704), *p*=0.031)

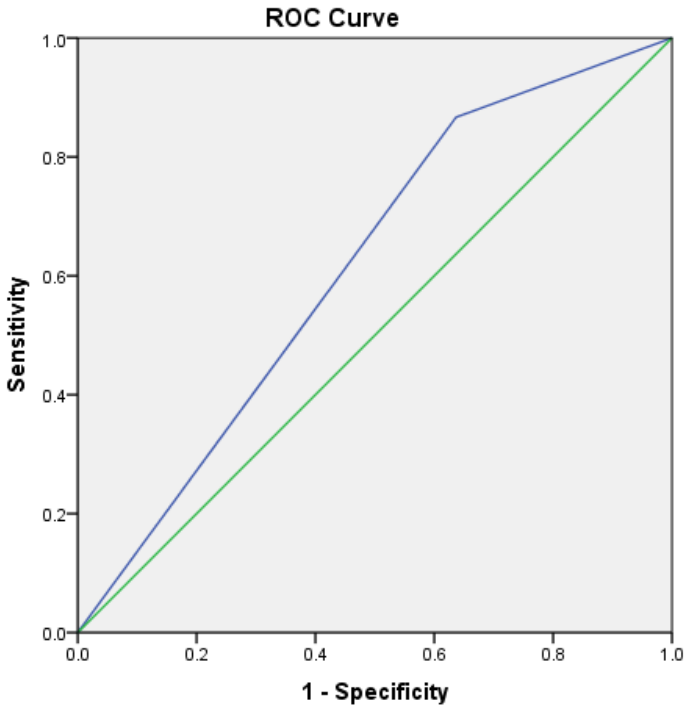
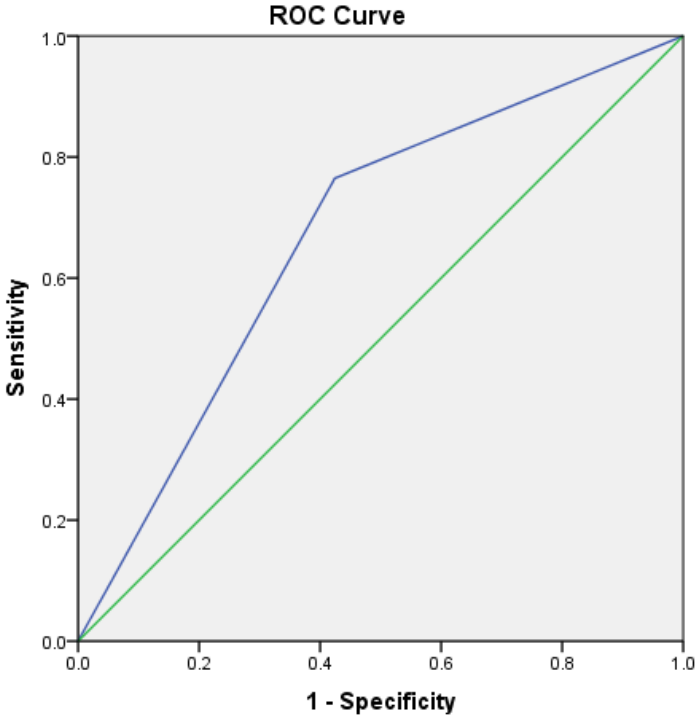


Figure 8. ROC for invasive cancer diagnosis over total patient follow up time, comparing the NICE high risk group (*BRCA* carriers included) to the NICE low risk group (AUROC=0.670



4.2.9. Kaplan-Meier analysis

The reported *p*-values are the Log Rank (Mantel-Cox) *p*-value from KM analysis.

Low and moderate risk group

Across the entire period of follow up time, there were 13 invasive breast cancers in the moderate risk group, and 4 in the low risk group. KM survival analysis demonstrated a significant difference in breast cancer rates across this time period between the low and moderate risk group *(p*=0.048). The KM survival curve is shown in *Figure 9*.

There was only 1 case of invasive breast cancer at <39 years in the moderate risk group. There were no cases of breast cancer <39 in the low risk group. The KM analysis was not significant (*p*=0.341).

Between the ages of 40-49, there were 4 invasive cancers in the moderate and 2 in the low risk group. There was no significant difference in the cancer rates across this age period between the low and moderate risk group (*p*=0.431).

From age 50-59 years, there were 8 invasive cancers in the moderate risk group and 2 in the low risk group. KM analysis demonstrated a significant difference in rates of breast cancer between the low and moderate risk group across this time period (*p*=0.037). The survival curve is shown in *Figure 10*.

Figure 9. KM analysis of invasive breast cancer in the low and moderate NICE risk categories across total patient follow up time. Log-Rank *p*=0.048.

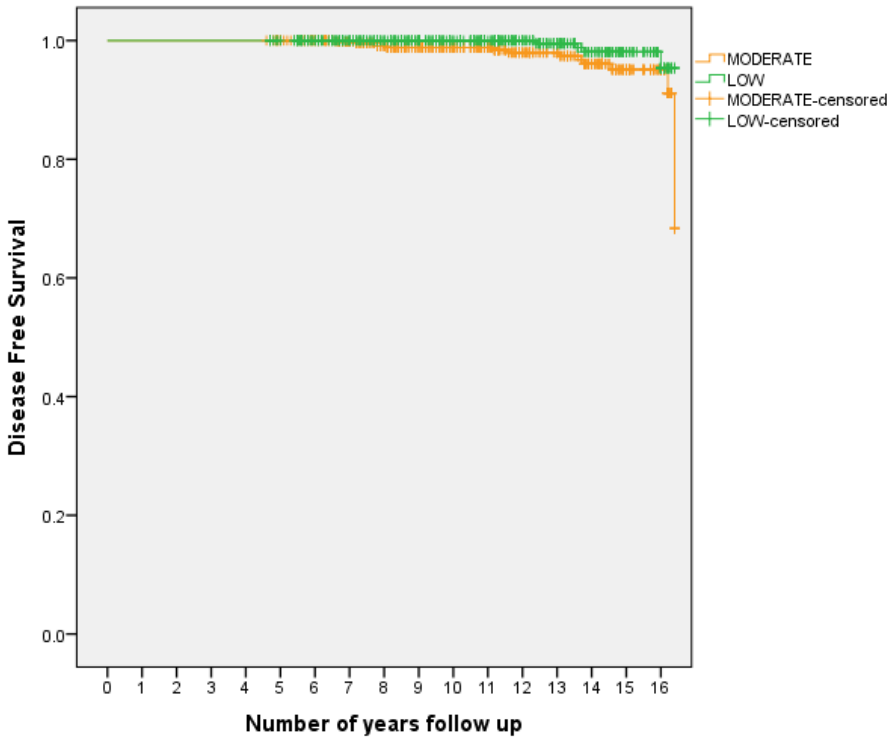
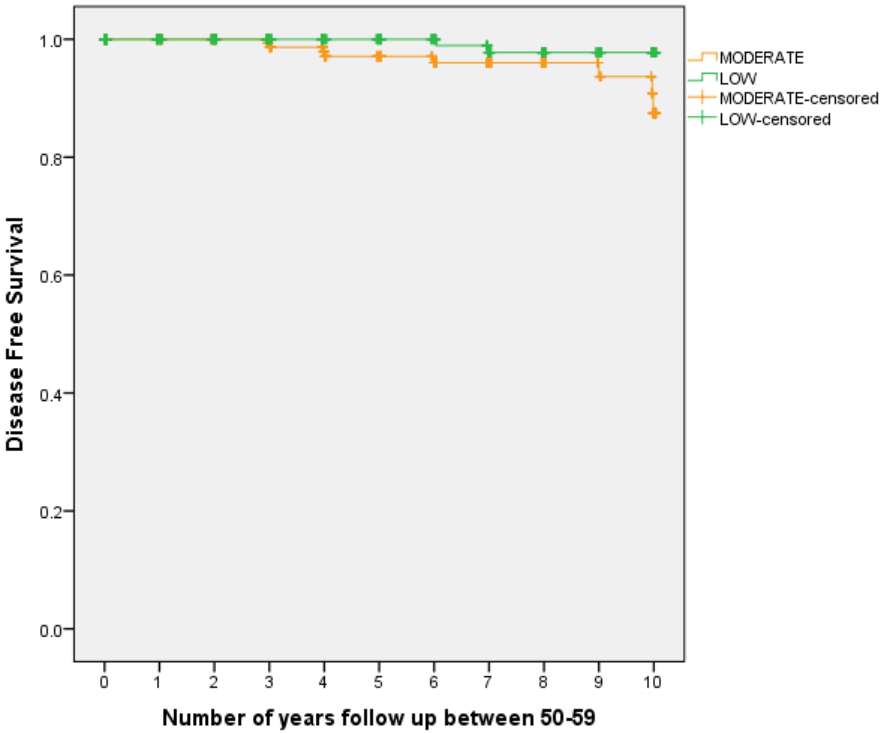


Figure 10. KM analysis of invasive breast cancer in the low and moderate NICE risk categories between ages 50-59. Log-Rank *p*=0.037



Low and high risk group (BRCA carriers included within the high risk group)

Across total patient follow up time, 13 invasive cancers occurred in the high risk group, and 4 in the low risk group. KM analysis demonstrated a statistically significant difference in rates of invasive cancer between the two groups (p=0.003). The survival curve can be seen in *Figure 11.*

There were 2 cases of invasive breast cancer at <39 years in the high risk group. There were no cases of breast cancer <39 in the low risk group. The KM analysis was not significant (p=0.091).

Between 40-49 years of age, there were 6 cases of invasive breast cancer in the high risk group, and 2 cases of breast cancer in the low risk group. KM analysis showed a significant difference in breast cancer rates between the low and high risk groups, between ages 40-49 (p=0.036). The survival curve can be seen in *Figure 12.*

From 50-59 years of age, there were 2 invasive cancers in the low risk group and 5 invasive cancers in the high risk group. On KM analysis, there was no significant difference in rates of invasive breast cancer development during this age range (p=0.149).

Figure 11. KM analysis of invasive breast cancer in the low and high NICE risk categories (*BRCA* carriers included) across total patient follow up time. Log-Rank *p*=0.003.

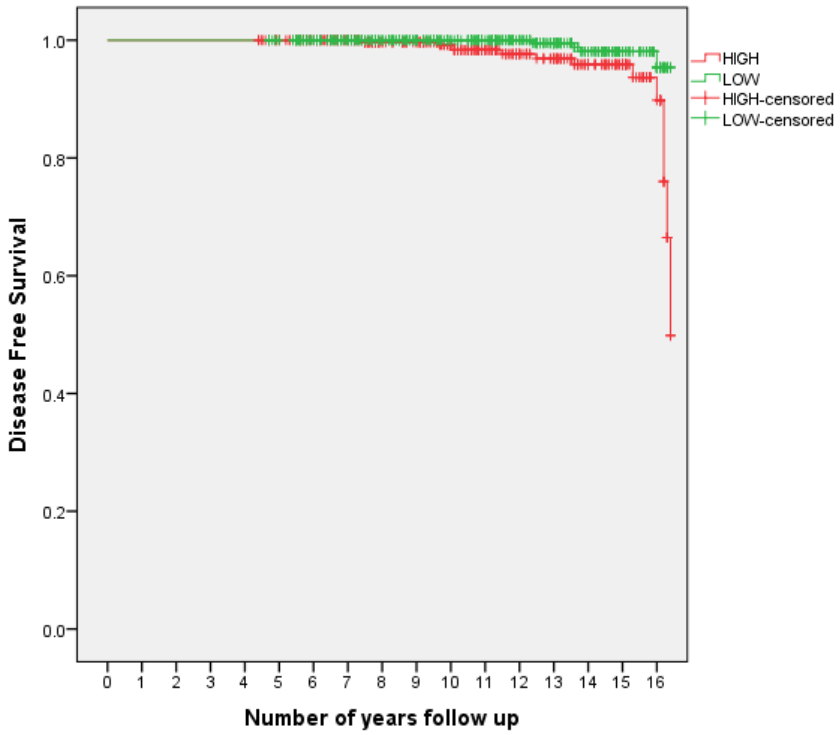
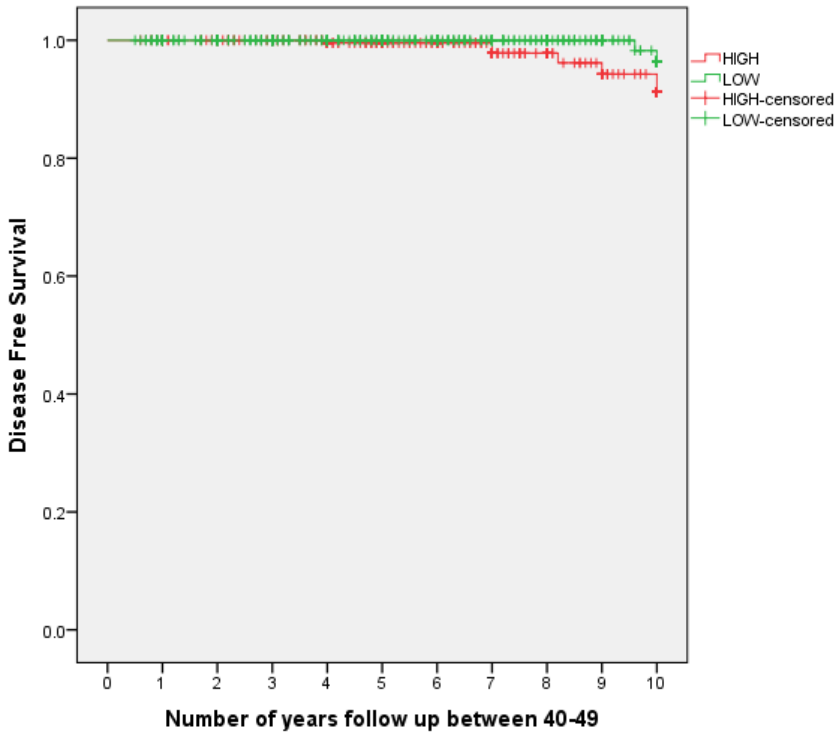


Figure 12. KM analysis of invasive breast cancer in the low and high NICE risk categories (*BRCA* carriers included) between ages 40-49 years. Log-Rank *p*=0.036.



Low and high risk group (BRCA carriers excluded from high risk group)

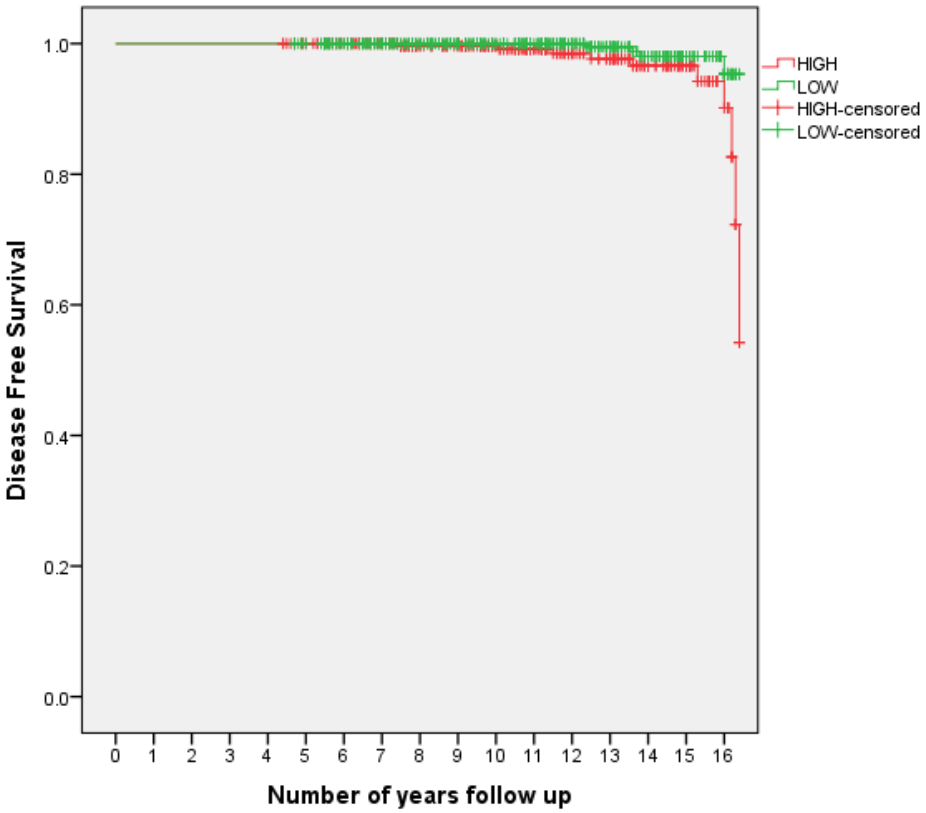
Excluding *BRCA* mutation carriers from the analysis, across total patient follow up time, there were 10 invasive cancers in the high risk group, and 4 invasive cancers in the low risk group. On analysis of rates of invasive breast cancer between the two groups, there was a significant difference across total patient follow up time (p=0.019). The KM survival curve can be seen in *Figure 13.*

At <39 years of age, there were 2 invasive breast cancers in the high risk group and 0 in the low risk group. The difference in invasive breast cancer rates over this time period between the two groups did not reach statistical significance (p=0.085).

Between the ages of 40-49, there were 4 invasive cancers in the high risk group, and 2 in the low risk group. The KM analysis demonstrated no significant difference in the rates of breast cancer between the low and high risk group between age 40-49, when *BRCA* carriers were excluded from the analysis (p=0.136).

There were 4 invasive breast cancers in the high risk group, and 2 in the low risk group between the ages of 50-59. KM analysis did not show a significant difference in rates of breast cancer during this age range between the high risk (excluded *BRCA* carriers) and low risk group (p=0.145).

Figure 13. KM analysis of invasive breast cancer in the low and high NICE risk categories (*BRCA* carriers excluded) across total patient follow up time. Log-Rank *p*=0.019.



Moderate and high risk group (BRCA carriers included)

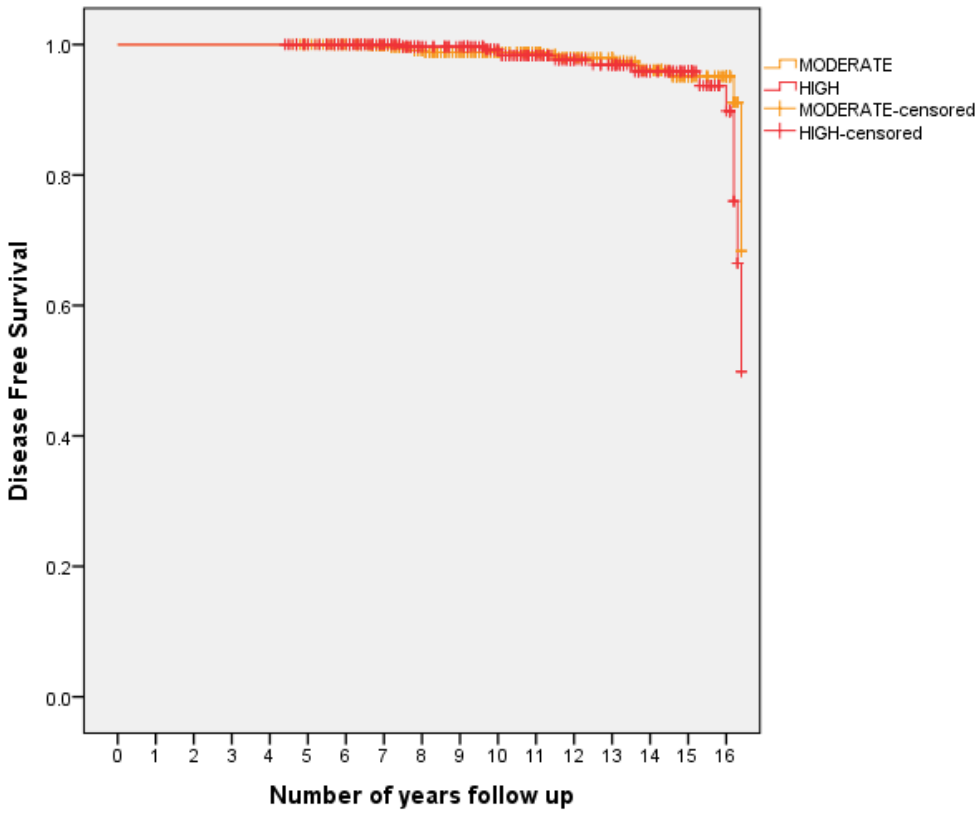
Across total patient follow up time, there were 13 invasive breast cancers in both the high and moderate risk group. KM analysis demonstrated no significant difference in invasive breast cancer rates across total patient follow up time between the moderate and high risk groups (p=0.274) The KM survival curve can be seen in *Figure 14.*

Aged <39 years there were 2 invasive cancers in the high risk group, and 1 in the moderate risk group. There was however, no statistically significant difference in the breast cancer rates across this time period between these two group on KM analysis (p=0.328).

There were 6 invasive breast cancers in the high risk group, and 4 in the moderate risk group diagnosed between the ages of 40-49. KM analysis demonstrated no significant difference in invasive breast cancer diagnosis during this time period between the moderate and high risk group (p=0.183).

From the ages of 50-59, there were 5 invasive cancers in the high risk group and 8 in the low risk group. There was no statistically significant difference in breast cancer rates between ages 50-59 comparing these two groups on KM analysis (p=0.581).

Figure 14. KM analysis of invasive breast cancer in the moderate and high NICE risk categories (*BRCA* carriers included) across total patient follow up time. Log-Rank *p*=0.274.



Moderate and high risk group (BRCA carriers excluded)

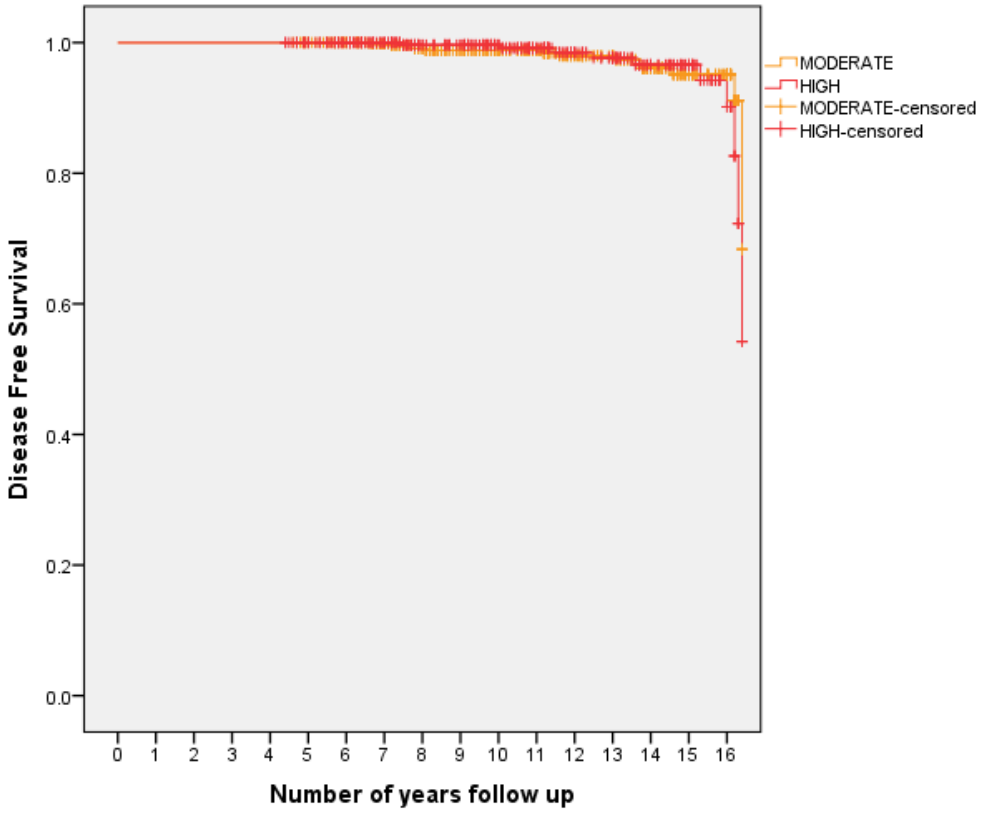
Excluding *BRCA* mutation carriers there were 10 invasive breast cancers in the high risk group, and 13 invasive breast cancers in the moderate risk group, throughout total patient follow up time. Comparing rates of invasive breast cancer diagnosis using KM analysis, demonstrates no significant difference between the moderate and high risk groups across total patient follow up time (p=0.644). The survival curve showing the two groups can be seen in *Figure 15.*

At age <39 years, there were 2 cancers in high risk group, and 1 in the moderate risk group. There was no statistically significant difference in breast cancer rates across this time period between the two groups (p=0.299)

From the ages of 40-49 there were 4 invasive breast cancers in the high risk and 4 in the moderate risk group. KM analysis demonstrated no significant difference between these two groups breast cancer rates from age 40-49 (p=0.499).

Between ages 50-59 years there were 4 and 8 invasive breast cancers in the high and moderate risk groups respectively. Again, there was no significant difference in breast cancer rates between these two group from the age of 50-59 using KM analysis (p=0.598).

Figure 15. KM analysis of invasive breast cancer in the moderate and high NICE risk categories (*BRCA* carriers excluded) across total patient follow up time. Log-Rank *p*=0.644.



Low and combined moderate and high risk group (BRCA carriers included)

An analysis was performed to examine whether there was a significant difference in rates of invasive breast cancer, between the low risk group, and the groups with any increase in risk of breast cancer (i.e. combined moderate and high risk). Across total patient follow up time, there were 26 cancers which developed in the combined moderate/high risk group, and 4 in the low risk group. KM analysis demonstrated a significant difference in breast cancer rates across total follow up time (p=0.011). The KM survival curve is shown in *Figure 16.*

There were no breast cancers in the low risk group at age <39 years and 3 breast cancers in women with any increase in risk according to NICE. However, there was no statistically significant difference in breast cancer rates using KM analysis between the two groups at age <39 years (p=0.216).

Between the ages of 40-49, there were 2 breast cancers in the low risk group and 10 in women of any increased NICE risk category. However, KM analysis showed no significant difference in rates of breast cancer across this time period (p=0.134).

There were 2 invasive breast cancers in the low risk group between the ages of 50-59 years and 13 invasive breast cancers in women of any increased NICE risk category over the same age range. There was found to be a borderline significant difference in breast cancer rates comparing these groups from age 50-59, using KM analysis (p=0.050). The survival curve can be seen in *Figure 17.*

Figure 16. KM analysis of invasive breast cancer in the low and moderate or high NICE risk categories (*BRCA* carriers included) across total patient follow up time. Log-Rank *p*=0.011

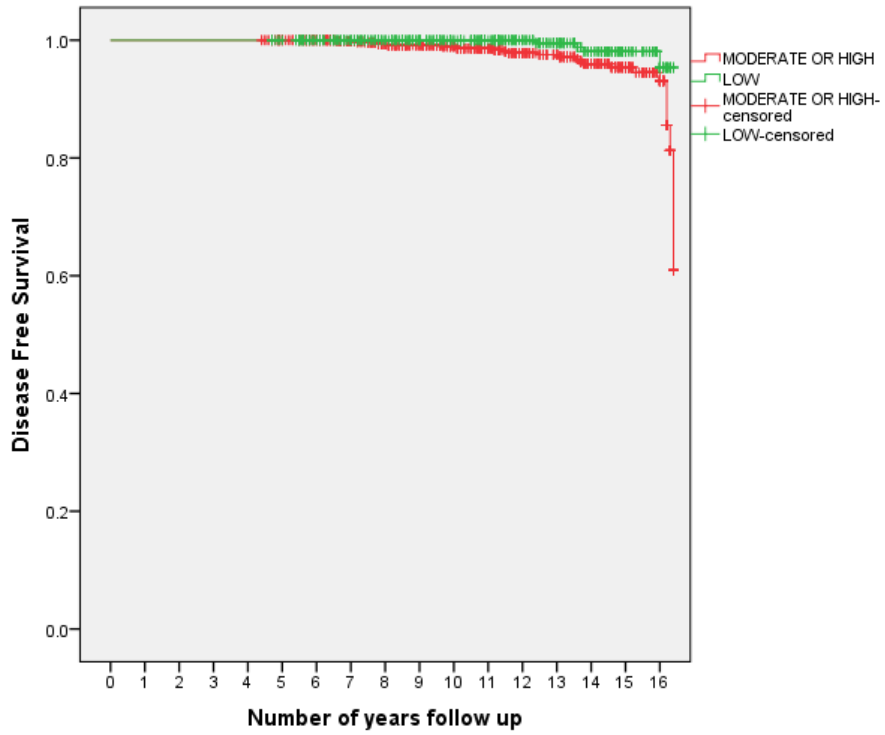
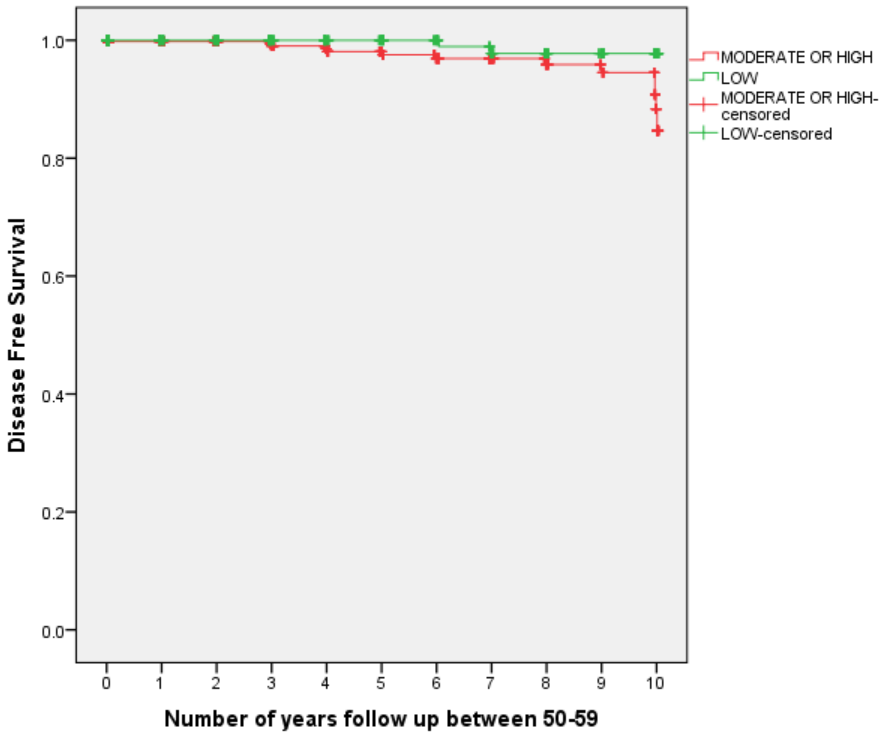


Figure 17. KM analysis of invasive breast cancer in the low and moderate or high NICE risk categories (*BRCA* carriers included) from age 50-59. Log-Rank *p*=0.050.



Low and combined moderate and high risk group (BRCA carriers excluded)

Another analysis was performed combining the moderate and high risk groups but with *BRCA* carriers excluded. Across total patient follow up time, there were 4 and 23 invasive cancer diagnoses in the low and moderate/high risk groups respectively. On KM analysis, this difference in rates of invasive breast cancer was found to be significantly different (p=0.024). The KM survival curve can be seen in *Figure 18*.

There were 3 cases of breast cancer in the moderate/high risk group at age <39, and none in the low risk group. However, this difference did not reach statistical significance using KM analysis (p=0.217).

From the ages of 40-49 years, there were 8 invasive breast cancer diagnoses in the moderate/high risk group, and 2 in the low risk group. Again, KM analysis did not reveal a statistically significant difference in rates of invasive breast cancer comparing the groups across the 40-49 age range (p=0.241).

Between the ages of 50-59, there were 12 invasive cancer diagnoses in the moderate/high risk group, and 2 in the low risk group. On KM analysis comparing rates of breast cancer diagnoses during this time period, there was found to be a modest statistically significant difference between the low and moderate/high risk groups (p=0.049). The KM survival curve is shown in *Figure 19.*

Figure 18. KM analysis of invasive breast cancer in the low and moderate or high NICE risk categories (*BRCA* carriers excluded) across total patient follow up time. Log-Rank *p*=0.024

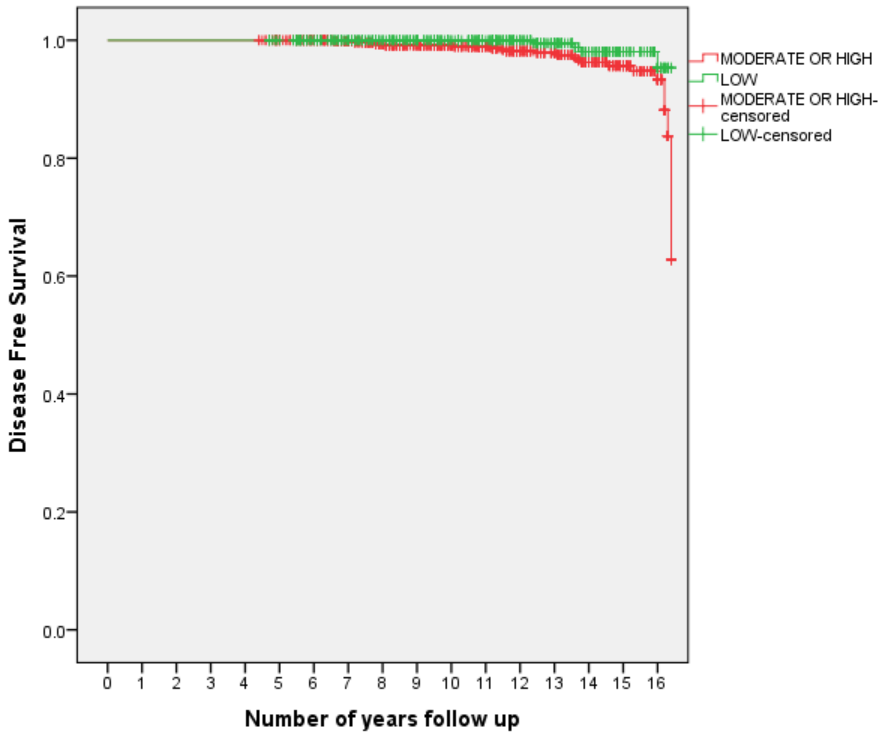
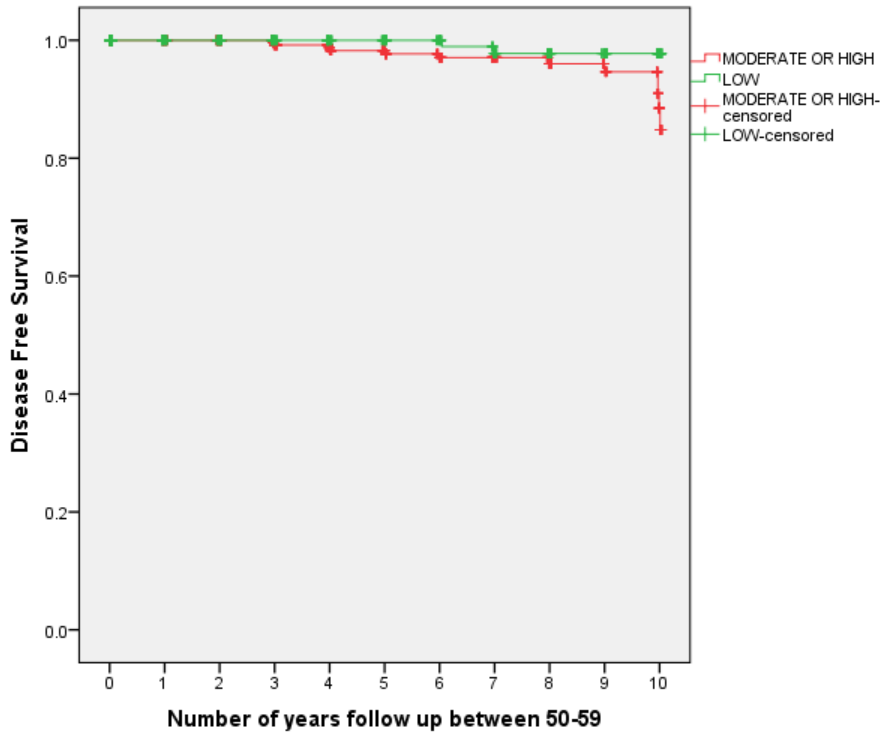


Figure 19. KM analysis of invasive breast cancer in the low and moderate or high NICE risk categories (*BRCA* carriers excluded) from age 50-59 years. Log-Rank *p*=0.049.



4.2.10. Summary of Kaplan-Meier analysis for invasive breast cancer

Table 29 shows a summary of the results of KM analysis comparing rates of invasive breast cancer diagnoses between different NICE risk categories across different age ranges. *P*-values shown are KM Log-Rank (Mantel-Cox) *p*-values. Significant results are highlighted in bold. *Figure 20* and *Figure 21* show Kaplan-Meier survival curves across total follow up time for low, moderate and high risk groups both including and excluding *BRCA* carriers.

Table 29. Summary of KM survival results comparing rates of invasive breast cancer between different NICE risk categories.

NICE risk groups being compared	KM Log-Rank (*p*-value)			
	Period of follow up			
	Total follow up time	Age ≤39 years	Age 40-49 years	Age 50-59 years
Low & moderate	**0.048**	0.341	0.431	**0.037**
Low & high	**0.003**	0.091	**0.036**	0.149
Low & high (*BRCA* carriers excluded)	**0.019**	0.085	0.136	0.145
Moderate & high	0.274	0.328	0.183	0.581
Moderate & high (*BRCA* carriers excluded)	0.644	0.299	0.499	0.598
Low & moderate/high	**0.011**	0.216	0.134	**0.050**
Low & moderate/high (*BRCA* carriers excluded)	**0.024**	0.217	0.241	**0.049**

Figure 20. KM survival curve showing invasive breast cancer rates for the low, moderate and high NICE risk groups across total patient follow up time (*BRCA* carriers included)

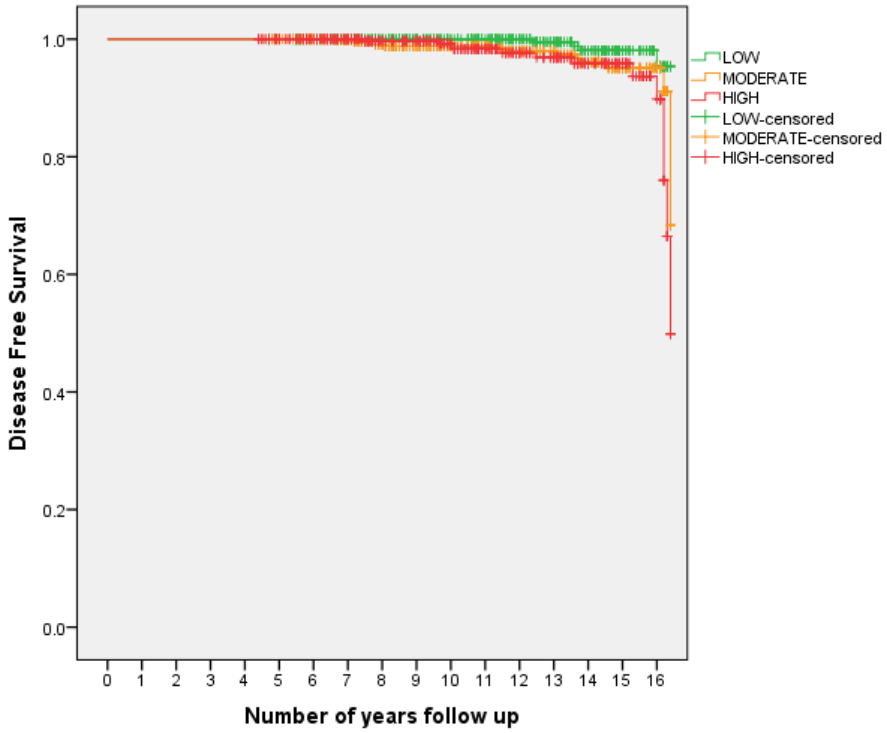
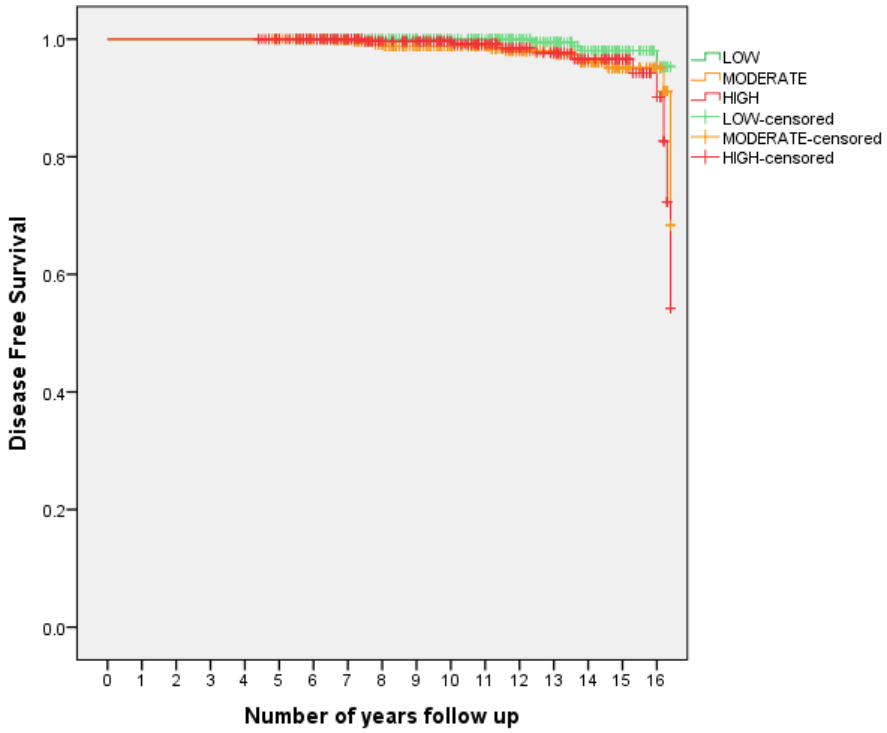


Figure 21. KM survival curve showing invasive breast cancer rates for the low, moderate and high NICE risk groups across total patient follow up time (*BRCA* carriers excluded)



4.2.11. Risk Summary

Table 30 shows summary information of the frequency and % 10-year absolute risk of invasive breast cancer, KM analysis and RR for each NICE risk group

Table 30. Summary data for NICE risk categories A) *BRCA* mutation carriers analysed separately B) *BRCA* mutation carriers analysed within high risk group

A)	**Number of women**	**Number of invasive cancers (% 10-year absolute risk (95% CI))**				**KM Log-Rank[a] *p*-value**				**Overall RR[a] (95% CI)**
		Overall	<39	40-49	50-59	Overall	<39	40-49	50-59	
Low	505	4	0	2 (0.82% (0.72-0.94))	2 (1.61% (1.42-1.83))	-	-	-	-	-
Moderate	522	13	1	4 (1.68% (1.53-1.83))	8 (7.05% (6.78-7.31))	0.048	0.341	0.431	0.037	3.14 (1.03-9.58)
High	360	10	2	4 (2.49% (2.28-2.70))	4 (5.28% (4.93-5.64))	0.019	0.085	0.136	0.145	3.51 (1.11-11.10)
***BRCA* carrier**	22	3	0	2 (26.67% (17.98-37.63))	1 (52.63% (31.71-72.67))	-	-	-	-	17.22 (1.10-72.29)
Total	1409	30	3	12	15					
B)	**Number of women**	**Number of invasive cancers (% 10-year absolute risk (95% CI))**				**KM Log-Rank[a] *p*-value**				**Overall RR[a] (95% CI)**
		Overall	<39	40-49	50-59	Overall	<39	40-49	50-59	
Low	505	4	0	2 (0.82% (0.72-0.94))	2 (1.61% (1.42-1.83))	-	-	-	-	-
Moderate	522	13	1	4 (1.68% (1.53-1.83))	8 (7.05% (6.78-7.31))	0.048	0.341	0.431	0.037	3.14 (1.03-9.58)
High	382	13	2	6 (3.56% (3.34-3.80))	5 (6.44% (6.10-6.78))	0.003	0.091	0.036	0.149	4.30 (1.41-13.07)
Total	1409	30	3	12	15					

4.2.12. Comparison with all women in the Tayside population who developed breast cancer age <50 years

Proportion of women assessed in clinical genetics who developed breast cancer

From 2000-2016, there were a total of 1,074 cases of female breast cancer diagnosed under the age of 50 in Tayside. This includes 15 of the 30 cancers identified from women in the original cohort. 267 of the women were seen in genetics at some time. However, only 47 were seen in clinical genetics prior to their breast cancer diagnosis, regarding breast cancer family history (4 were seen regarding another genetic condition). 220 were seen after their breast cancer diagnosis. This leaves 807 women in Tayside who developed breast cancer under 50, from 2000-2016, and have to date never received any genetic counselling. In total, there were 1,027 women who developed breast cancer at age less than 50 years from 2000-2016 who had not previously been seen in genetics therefore had likely not received any additional screening. This information is summarised in *Table 31.* The proportion of women in Tayside with breast cancer under the age of 50 for the given time period, who were previously assessed by the clinical genetics department is therefore 4.38% (95% CI, 3.31-5.78%).

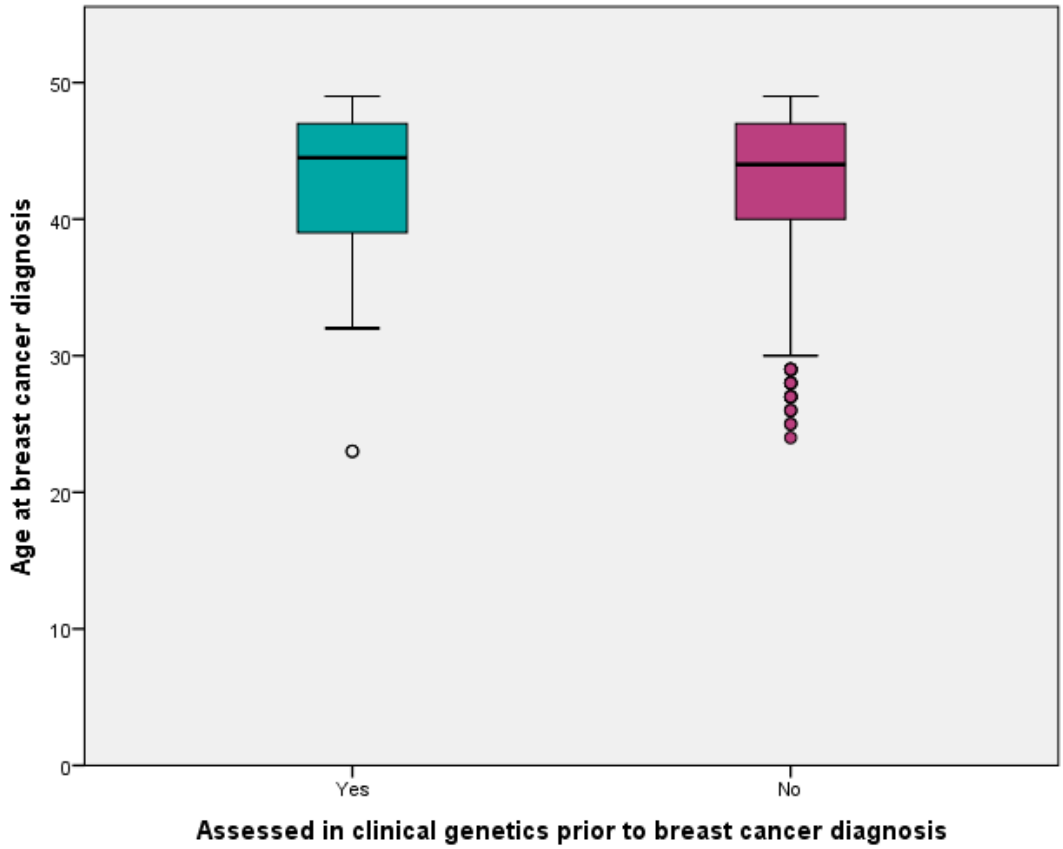
Table 31. Breakdown of women who were/were not seen in clinical genetics who developed breast cancer under the age of 50 in Tayside from 2000-2010

<table>
<tr><td rowspan="6">Assessed in clinical genetics</td><td rowspan="5">Prior to breast cancer diagnosis</td><td>Low risk</td><td>15</td><td rowspan="6">267</td></tr>
<tr><td>Moderate risk</td><td>7</td></tr>
<tr><td>High risk</td><td>21</td></tr>
<tr><td>BRCA carrier</td><td>4</td></tr>
<tr><td>Total</td><td>47</td></tr>
<tr><td colspan="2">After breast cancer diagnosis</td><td>220</td></tr>
<tr><td colspan="3">Never assessed in clinical genetics</td><td colspan="2">807</td></tr>
<tr><td colspan="3">Total</td><td colspan="2">1,074</td></tr>
</table>

Age at diagnosis comparison

The mean age at diagnosis of the women seen in clinical genetics, prior to their breast cancer diagnosis, was 42.46 years (SD±6.112). The figure for women not seen prior to their breast cancer diagnosis was 42.85 years (±5.139). On Independent T-test analysis, there was no significant difference between the ages at breast cancer diagnosis between the two groups (*p*=0.598). The boxplot for age at diagnosis is shown in *Figure 22*.

Figure 16. Boxplot comparing age at breast cancer diagnosis between women seen in clinical genetics prior to diagnosis and women not seen prior (p=0.598)



5.0. Discussion

5.1. Cohort characteristics

The initial data collection identified 1,409 patients for inclusion in the main analysis of this research, with a total of 15,414 patient years of follow up. The mean number of years follow up per woman was 10.935 years (SD±3.3311). Women within the cohort had presented to clinical genetics at an average age of 36.75 (±7.855) years, and women who presented over the age of 50 were not included. The family structures and cancer history of patients included varied, however the most common relative reported to be affected with breast cancer was the patient's mother (60.6%). The mean number of relatives reported to be affected with breast cancer was 1.649, however some families had up to 7 affected FDRs or SDRs. The average age of a relative at breast cancer diagnosis was 50.308 (±10.817) years. Overall, it is likely that the data reflects a standard patient who would be attending clinical genetics regarding a FH of breast cancer, and 15, 414 patient years of follow up provides a considerable quantity of data for analysis.

5.2. NICE risk category

5.2.1. NICE risk category assignment

505 (35.84%) of the women were assigned a low risk category, 522 (37.04%) moderate risk, and 382 (27.11%) high risk, suggesting a reasonable distribution of women in each group. There was no significant variability in the age of presentation between the groups (p=0.092), therefore reducing the likelihood that the follow up time at certain age ranges was significantly different. It is worth noting that, all-together, 64.1% of women who attended clinical genetics about their breast cancer risks were eligible to be assigned to a category of increased risk. The majority of women referred are therefore appropriate for additional screening. Most of the women in the moderate risk group were assigned that risk due to a FH of 2 FDRs or SDRs diagnosed with breast cancer at an average age of older than 50 years (58.3%). For the high risk women, the most common reason was a FH of 2 FDRs or SDRs, diagnosed with breast cancer at an average age of younger than 50 years (32.1%).

This suggests that the age of relatives at diagnosis is a key determinant of which risk category women are assigned according to NICE guidance, especially given that the mean number of relatives affected in the cohort was 1.649.

The family histories used to categorise patients were correct according to clinical records, and most had been confirmed by the relevant cancer registry. Studies on the reliability of FH reporting have demonstrated that self-reported familial cancer history tends to be accurate, with over-reporting a rare event (225, 226). Therefore, there is reasonable confidence that the family histories used to categorise women are accurate, and the results here reflect the risks in this cohort.

5.2.2. Potential discrepancies in NICE risk criteria

Whilst assigning NICE risk categories, some potential pitfalls in the criteria were identified. The relevance of this clinically is unclear given that, for the purposes of this study, the NICE guidelines were interpreted in a literal sense. For example, other than women with four or more affected family members, the guidance refers to only FDRs and SDRs affected with breast cancer, and doesn't account for more distant relatives. It could be argued that a relative more distant than a SDR, shares too little of their genetic material with the patient to be relevant to polygenic mode of inheritance. However, it is important to consider intervening male relatives. The estimated incidence of male breast cancer is <1 per 100,000 men per year, and even amongst male *BRCA* carriers the lifetime risk of breast cancer is around 1-5% for *BRCA1* and 5-10% for *BRCA2* (227). Therefore, if there is a familial predisposition to breast cancer segregating in a family, when linked by 1 or more successive males it may not be phenotypically expressed in close relatives of the female patient. For women at a moderate risk due to presumed polygenic effects, this may be less relevant, as it would be assumed that a polygenic risk of breast cancer would be diluted within a few generations. The MMSS takes into account intervening male relatives when scoring for a >10% likelihood of a *BRCA* mutation in the family, by moving female relatives connected by a male up one degree of relation. It therefore bypasses this issue. However, this is not expressly stated in the NICE guidelines.

In addition, there were some cases of women who had a FH of breast and ovarian cancer, however there were not enough breast cancers, or they were not diagnosed young enough to meet the high risk criteria. These women, for the purposes of the study, had to be assigned to the low risk group. The same applied for women with a FH of male breast cancer. The female breast cancers required to meet high risk criteria for those with a FH of ovarian or male breast cancer must have been in an FDR or SDR. Male breast cancer and ovarian cancer, in the context of a FH of female breast cancer, may be associated with a *BRCA* mutation. Therefore, it would make sense that the breast cancers would have occurred at a younger age. However, this cannot be determined with absolute certainty. In addition, if a *BRCA* mutation were segregating through male relatives, there may be more than 1 unaffected male *BRCA* carrier connecting two female carriers – therefore the closest female breast cancer may be more distant than a SDR.

The theoretical pitfalls described above explain why some women in the low risk group had a very high MMS despite being low risk according to NICE. Hypothetically, if a women was low risk according to NICE, it is possible that a MMS would not be calculated for them at all. This means that, in theory, women who met the >10% mutation risk threshold may be missed. In reality, clinical judgement will determine what to do when faced with anomalies such as these. The impact of these considerations on clinical management is therefore uncertain in the context of this research. These issues leave room for varied interpretation of guidance amongst different clinicians, and between different centres in the NHS. Patients in different areas may therefore receive different care, despite having a similar FH. Whilst clinicians should always be interpreting guidance with their own knowledge in mind, it would be desirable to have guidance which avoids elements of uncertainty and of which there can be universal interpretation.

5.3. The Modified Manchester Score tends to increase with NICE risk category and predicts *BRCA* mutations in the cohort

MMS was calculated for each patient in the cohort and compared between NICE risk categories. The mean MMS for the low, moderate and high risk groups was 6.57 (SD±5.01), 9.03 (±3.76) and 17.61 (±8.68) respectively. Despite this there were still some significant outliers in the low and moderate risk group due to factors discussed in section 5.2.2. 15.6% of the cohort (220 patients) had a MMS ≥17 however only 62 patients in the cohort had in fact been tested for a *BRCA* mutation. Of interest, only 33 of these patients themselves had a MMS of 17 or above. Of the remaining patients who had been tested, all but one had a known *BRCA* mutation segregating in the family. 22 of the 62 tested were *BRCA* mutation positive.

Analysis was performed to assess whether the cut-off MMS of 17 predicted a *BRCA* mutation. This result of Fischer's exact test (*BRCA1* (*p*=<0.001)*, BRCA2* (*p*=0.002) or either mutation (*p*=<0.001)), and AUROC analysis ((*BRCA1* (*p*=<0.001)*, BRCA2* (*p*=0.015) or either mutation (*p*=<0.001)), were significant. This suggests benefit of the ICR MMSS (210) in helping identify women who would benefit from mutation testing. This is despite a relatively small sample size and the fact that 7 of the 22 *BRCA* carriers had a MMS less than 17. The MMS was generally more specific than it was sensitive. However without the mutation status of all the women with a score of 17 or greater this cannot be quantified with certainty, therefore it is difficult to interpret much from these results. Despite this, the MMS would appear to have some clinical utility. It would be beneficial to assess this in a larger group of women tested for a *BRCA* mutation, given the limited number of known mutation carriers in this study.

5.4. Independent T-test and Pearson Chi-square/Fischer's exact analysis find few predictors of breast cancer development in family history

To tease out which elements of a patient's FH may correlate significantly with their overall and age specific risk of breast cancer, analysis was performed to assess different FH factors. Risk of cancer across all follow up time, was found to significantly correlate with the number of FDRs and SDRs in total affected in the family, both including (p=0.004) and excluding (p=0.003) *BRCA* carriers from analysis. The number of FDRs affected was found to be significant when including *BRCA* carriers in analysis (p=0.024), but not when excluding them. Having ≥2 affected FDRs or SDRs was significantly associated with overall breast cancer risk when *BRCA* carriers were excluded or included (p=0.003 for both). However, there was no correlation with any of the other variables assessed, including the age of relatives at diagnosis (and specific age cut-offs), or FH of ovarian cancer, bilateral breast cancer or male breast cancer. This would suggest that, overall, having 2 or more FDRs or SDRs affected is a predictor of developing breast cancer, but that in this cohort, age of relatives at diagnosis, or any other factor did not seem to affect the risk.

For cancer diagnoses under the age of 39 years, none of the variables assessed were determined to be significant both excluding and including *BRCA* mutation carriers. However, in the whole cohort, only 3 women developed breast cancer at age <39 years. It is therefore difficult to interpret the relevance of these results. Breast cancer at age <39 years is uncommon, affecting just 92/100,000 women in the UK population (1). Therefore, even in a population such as this which is enriched for women at an increased risk of breast cancer, a much larger cohort size would be required to detect enough women for sensible statistical analysis.

Decreasing average age at diagnosis of FDRs significantly correlated with breast cancer development between the ages of 40-49 years, when *BRCA* carriers were excluded from analysis. Otherwise, none of the variables correlated significantly with breast cancer in this age range, including number of FDR/SDRs. Conversely, the only predictor of breast cancer between age 50-59 years appeared to be the total number of FDRs and SDRs affected (*BRCA* excluded p=0.038, *BRCA*

included *p*=0.036), however there was no significant cut-off above which a certain number of affected relatives became significant.

It is surprising that decreasing age of relatives at diagnosis significantly correlates with breast cancer between ages 40-49 when excluding, but not including *BRCA* mutation carriers. Since families with *BRCA* mutations demonstrate young onset breast cancers, the significance would be presumed to increase by including them in the analysis. This may be due to a lack of young onset breast cancers in the female FDR or SDRs of *BRCA* carriers, given that it may also present with male breast cancers and ovarian cancers. Again, since the MMS takes into account intervening males, women may be selected for *BRCA* testing despite a relative lack of affected close female relatives. Alternatively, and perhaps more likely, this may be a spurious result given the relatively small number of *BRCA* carriers within the cohort. Supporting this theory, ovarian, bilateral and male breast cancer did not reach significance when *BRCA* carriers were included despite being well established features of families with *BRCA* mutations.

It is possible that younger age of relatives at diagnosis may be more relevant for individuals who will go on to develop breast cancer at <50 years of age, rather than the total number of affected relatives. The total number of relatives even when diagnosed at older ages may be more relevant in predicting the risk of breast cancer over the age of 50, as the data here would suggest. Given that there were relatively small numbers of cancers in each age group, the most relevant results to interpret are likely to be the results pertaining to the whole cohort, over all age ranges. These results suggest that the total number of FDRs and SDRs diagnosed is the strongest predictor of developing breast cancer, with ≥2 relatives diagnosed linked to a significant risk. Overall, age of relatives at diagnosis doesn't appear to be relevant, other than perhaps in the 40-49 age group. Even then, cut offs of an average age of diagnosis at <40 years (*BRCA* excluded *p*=0.856, *BRCA* included *p*=0.996) or <50 years (*BRCA* excluded *p*=0.083, *BRCA* included *p*=0.275) seemed to make no difference to breast cancer risk. This is in disagreement with previous literature; an analysis of family history and breast cancer risk from the Nurses' Health Study (151) found no significant difference in cancer risk for women with a mother (*p*=0.06) or sister

(*p*=0.43) affected with breast cancer under the age of 50, compared to over the age of 50. However, when analysed cumulatively, women with either a mother or sister (i.e. any female FDR) diagnosed under the age of 50, had a significantly higher risk of breast cancer than women with a FDR diagnosed over 50 (RR=1.70 and 1.30 respectively, *p*=0.016). The authors concluded that this was a useful way to stratify risk. However, the associated breast cancer risk for the women in the cohort was not age stratified, therefore whether the risk applies to young onset breast cancers is unclear.

As described in section 5.2.1, the most common reasons for being assigned to either the NICE moderate or high risk category can be differentiated by the average age of diagnosis of FDRs or SDRs. The cut-off is 50 years of age. In this cohort, a cut-off of <50 years average age at diagnosis did not appear to be significantly correlated with risk of breast cancer at any age range. This suggests that a cut-off of less than, or greater than age 50 at average breast cancer diagnosis, has limited use in distinguishing moderate from high risk women, despite being a differentiator for the most common presenting family histories.

5.5. Overall mean age of cancer development is not significantly different between risk groups

The mean age of women at breast cancer diagnosis for each risk group was calculated and analysed using one-way ANOVA. Although there was a clear trend towards younger age at diagnosis with increasing risk category, as demonstrated in *Figure 6* (section 4.2.1.), this did not reach significance (p=0.578). This trend may be expected if breast cancers are picked up earlier due to inclusion in early screening. In addition, this is not necessarily the best analysis of age at cancer diagnosis between the groups, as it is a static test. It does not take into account the age of each patient at initial consultation or their overall follow up time. This is important since women who had previously been diagnosed with young onset breast cancer were not included in the study. Therefore, partly as a product of the study design, there are few breast cancer diagnoses under the age of 40. A time-dependent analysis is therefore more beneficial.

5.6. Cancer risk for each NICE category in this cohort does not meet the absolute percentage 10-year risk suggested by guidelines

As shown in *Table 25* (section 4.2.5.), there were 4 cancers in the low risk group, 13 in the moderate risk group, and 13 in the high risk group, of which 3 were accounted for by *BRCA* mutation carriers. 15 of the cancers occurred between 50-59 years, 12 between 40-49 years and 3 cancers occurred at age <39 years. There were no cancers recorded at age 60 or above, however this is expected due to the younger age of women included in the study (<50 years) and the maximum follow up period of 16 years.

Using the number of women years of follow up calculated for each category between different age ranges (*Table 20,* section 4.1.6.), % absolute risk per woman year of follow up within the age ranges 40-49 and 50-59, was used to determine the approximate 10-year % absolute risk in that time period. This was to establish whether or not, in this cohort of women, the % 10-year absolute risk was approximately equal to that which NICE suggests their risk should equate to. NICE states that the % absolute risk between the age of 40-49 should equate to <3%, 3-8%, and >8% for the low, moderate and high risk groups respectively. It is at these risk levels which they deem the relevant additional screening to be appropriate.

In this cohort, the % 10-year absolute risk from age 40-49 for each category was 0.82% (95% CI, 0.72-0.94%) for the low risk group, 1.68% (1.53-1.83%) for the moderate risk group and 3.56% (3.34-3.80%) for the high risk group (when *BRCA* mutation carriers were excluded this risk was 2.49% (2.28-2.70%)). The risk of breast cancer in the cohort between the ages of 40-49, increased with increasing NICE risk category, suggesting some ability of the guidelines to differentiate risk groups. However, the risks and 95% CIs are substantially less than what NICE suggests is the appropriate cut off for additional screening. This would suggest that between the ages of 40-49, NICE guidance overestimates the risk of breast cancer associated with specific FH criteria.

Between the ages of 50-59 years, the absolute % 10-year risk for low, moderate and high risk women was 1.61% (95% CI, 1.42-1.83%), 7.05% (6.78-7.31%)

and 6.44% (6.10-6.78%) respectively (5.28% (4.93-5.64%) for high risk excluding *BRCA* carriers). Cumulatively, during the 20-year period from age 40-59, the high risk group was at the greatest % absolute risk in this cohort. However, it is interesting to note that for the age range of 50-59 the highest % risk occurs in the moderate risk group, as it reaches 7.05%. This is the highest risk for any single % 10-year age range. In addition, despite the risk being in the ages 50-59 range, rather than the 40-49 year age range as it should be, it is the only % absolute 10-year risk which crosses its level of clinical significance suggested by NICE (i.e. 3-8% for the moderate risk group).

The fact that the absolute risk for the moderate group appears greater than the high risk group between ages 50-59 is interesting. It is unlikely that the explanation is that women with a moderate risk FH are at a greater risk of breast cancer overall, given that the cumulative absolute risk from 40-59 is greater in the high risk group. A perhaps more likely explanation, is that it reflects women in the high risk group being diagnosed with cancer earlier, before the age of 50, and that the main proportion of risk for the moderate group occurs after the age of 50. Additionally, based on the data from this cohort, it would appear that there is a 3-8% risk for moderate risk women between 50-59, rather than 40-49 years. Women at moderate risk of breast cancer can be *considered* for annual screening from age 50-59 as part of the NICE guidelines, however they are offered it between ages 40-49. Based on this study's results, it would appear to be more beneficial for these women to be offered annual screening between ages 50-59 rather than just considered for it. The usefulness of screening moderate risk women at age <50 years is questionable given that in this cohort their absolute risk doesn't reach screening threshold. NICE guidance is formulated taking into account a cost model based on an 'incremental cost-effectiveness ratio' (ICER), which assesses cost, against quality of life years gained through an intervention. An ICER of <£20,000 is deemed cost effective, though there is no absolute cut off (228). Assuming that the familial breast cancer guidelines were formulated under this model, the fact that the absolute 10-year risk of breast cancer is overestimated, would suggest that the guidelines may not be cost effective.

5.7. Relative risks, odds ratios, sensitivity and specificity for each category

5.7.1. The relative risk and odds ratio for developing breast cancer increases with increasing NICE risk category

The RR and OR for each NICE category, as well as the combined moderate and high risk groups are shown in *Table 26* (section 4.2.6.). This was with comparison to the low risk group. The RR and OR for breast cancer development tended to increase with increasing NICE risk category. The highest risk, unsurprisingly, was for *BRCA* carriers with a RR of 17.22 (95% CI, 1.10-72.29) and an OR of 19.78 (4.13-94.63). The next greatest risk was for the high risk group including *BRCA* mutation carriers, (RR 4.30 (1.41-13.07), OR 4.410 (1.43-13.64)). As would be expected, this risk dropped when *BRCA* carriers were excluded from the high risk group (RR 3.51 (1.11-11.10), OR 3.58 (1.11-11.50)), however still remained greater than that of the moderate risk group (RR 3.14 (1.03-9.58), OR 3.20 (1.04-9.88)). It is worth noting the wide confidence intervals for these figures, due to the relatively small number of breast cancers that developed in the cohort, however the lower limit of the confidence interval is never less than or equal to 1.00.

5.7.2. The sensitivity of NICE risk categories is reasonable but the specificity is poor

The criteria for the moderate, high, or combined moderate/high risk group, had reasonably good sensitivity for invasive breast cancer development across the total period of follow up time (76.47% (95% CI, 49.8-92.17%), 76.47% (49.76-92.17%) and 86.67% (68.36-95.64%) respectively). In addition, the NPV was above 99% for the moderate, high or combined group, with narrow confidence intervals (see *Table 27,* section 4.2.7.), however this is expected given that the number of events was small. Specificity was 49.6% (95% CI, 46.5-52.7%) for moderate, 57.59% (54.22-6.89%) for high and dropped to 36.33% (33.80-38.94%) for either risk group. Likewise, the PPV for each group was never above 4%.

Although these values are not age specific, it is worth noting that the sensitivity was 86.67% for the combined risk group, as additional screening is offered from at least 40 for both. This means that theoretically, a considerable portion of

breast cancer which may occur in the cohort should be picked up through screening, although moderate risk women over the age of 50 do not necessarily have to be screened annually, and may return to the NBSP. Only 3 of the 30 cancers in the cohort occurred prior to age 40 (none of which were in the low risk group), and 2 occurred in the low risk group between ages 40-49. Therefore, in theory, 25 of 30 cancers (83.33%) which developed in the cohort had the potential to be picked up either by early screening due to increased NICE risk category, or through the NBSP. However, this seems to come at a considerable trade off with specificity. To be cost-effective, it is beneficial for screening programmes to be both as sensitive and specific as they can be. The sensitivity appears to be reasonably good, and a modest trade-off for specificity is generally acceptable if the benefits of screening are substantial. However, it is worth considering whether or not a specificity of maximally 57.59% in this cohort is cost-effective, especially given the small number of events. AUROC analysis, discussed below, provides a more informative measure of the effectiveness of the criteria at discriminating risk groups, combining sensitivity and specificity.

5.7.3. Using area under the receiver operating curve, NICE guidelines did not perform significantly better than chance for identifying women who would develop breast cancer when *BRCA* carriers are excluded

AUROC analysis was used to assess the performance of the risk categories overall, and at different age ranges, compared to the low risk group. The results are shown in *Table 28* (section 4.2.8.).

The high risk category performed significantly better as a breast cancer prediction model over all time periods (p=0.016), as did the combined moderate and high risk category (p=0.031), although the AUROC was only 0.670 and 0.615 respectively. However, the significance of the results was lost on exclusion of *BRCA* carriers. There was no significant difference in performance for specific age categories (<39 years, 40-49 years, 50-59 years). The moderate risk category did not perform significantly better than the low risk group at predicting breast cancer at any specific time period or overall.

Although the sensitivity was high for all increased risk categories, it would appear that the poor specificity in the moderate risk group limits its usefulness as a risk predictor. From the analysis of the RRs and ORs for each group, we can see that there is a definite trend towards increasing risk with risk category. However, the AUROC analysis would suggest that the risk is not significantly greater, with enough specificity in the moderate risk group for it to be an overall beneficial discriminator.

The loss of significance when excluding *BRCA* carriers from the high and combined risk group, also suggests that the performance of the risk categorisation criteria is dependent on the highest risk patients i.e. *BRCA* carriers. Excluding them from analysis, the model fails to effectively identify at-risk women. However, it is worth noting that the analysis, again, does not take into account the number of years follow up each woman in the cohort received, and that numbers in each subcategory were relatively small. Nonetheless, they will contribute to the overall assessment of the effectiveness of the guidelines. The most interesting analysis of risk, is the rate of breast cancer diagnosis across different age ranges, rather than the number which developed overall; it reflects the information most relevant to the question of when women at increased risk should be screened.

5.8. Kaplan-Meier Survival Analysis

5.8.1. The rate of breast cancer development is significantly greater in the moderate risk group than the low risk group, overall, and between age 50-59 years

KM analysis allows determination of the difference in follow-up time specific rates of diagnosis between groups, which can then be analysed by different age categories. When comparing the low and moderate risk group, overall, there was a significant difference in rates of breast cancer development between the moderate and low risk group (p=0.048). On analysis of specific age ranges, there was a significantly higher rate of breast cancer in the moderate risk group between the ages of 50-59 (p=0.037). However, from ages 40-49 and <39 years, there was no significant difference in rates between the two groups. This would suggest that women in the moderate risk group are at increased risk of breast cancer, but that the increased risk only becomes significant after the age of 50 years.

5.8.2. The rate of breast cancer development is significantly greater in the high risk group compared to the low risk group, overall, and between ages 40-49 years when *BRCA* carriers are included

The high risk group was analysed, both including and excluding *BRCA* carriers from the analysis. Across total follow up time, the high risk group demonstrated significantly increased rates of breast cancer development compared to the low risk group, including or excluding *BRCA* mutation carriers (p=0.003 and p=0.019 respectively). At <39 years, the difference in breast cancer rates between the high risk (including and excluding *BRCA* carriers) and low risk group was not significantly different, however there were very few cancers at this age. Likewise, from age 50-59 years the breast cancer rates for the high risk group were not significantly greater than the low risk group both including (p=0.149) and excluding (p=0.145) *BRCA* carriers.

The most interesting results on KM analysis in the high risk group were between 40-49 years. Including *BRCA* mutation carriers, the difference in breast cancer rates over this 10-year period was significant (p=0.036). However, when *BRCA*

carriers were excluded from the analysis, the value was no longer significant, at *p*=0.136. This would suggest that, other than for people with a *BRCA* mutation, the breast cancer rate prior to age 50 is not significantly different for women with a NICE high risk FH than for women with a NICE low risk FH. Likewise, the rate doesn't appear to be greater for them after age 50. Given that additional screening will be offered to these women, it is crucial to know whether or not the difference is significant.

5.8.3. There is no significant difference in breast cancer rates between the moderate and high risk group

KM analysis was also performed to assess the difference in breast cancer rates between the moderate and high risk group. Including or excluding *BRCA* carriers, there was no significant difference in rates of breast cancer diagnosis overall, or at any specific age range. It would appear that the NICE guidelines poorly differentiate between women at a moderate and women at a high risk of breast cancer. However, when these two categories are combined and breast cancer rates compared with the low risk group, we can see significant differences; including *BRCA* carriers and over total follow up time, the difference in rates is significant (*p*=0.011), and the same holds true when *BRCA* carriers are excluded (*p*=0.024). Therefore, although the differentiation between moderate and high risk is poor, overall, the NICE guidelines are able to differentiate between women at low risk and women at *some* increased risk of breast cancer. Analysing the combined moderate and high risk group also shows borderline significantly different breast cancer rates compared to the low risk groups between ages 50-59 years (including *BRCA* carriers *p*=0.050, excluding *BRCA* carriers *p*=0.049). There was no significant rate difference at <39 years or 40-49 years. Again, this is crucial as additional screening is offered to all in the combined risk group from age 40, despite no significant increase in risk detected in this study.

5.9. NICE guidelines identify women at some increased risk of breast cancer, mainly after the age of 50 years

The results of KM analysis provide an interesting insight into the effectiveness of NICE guidelines at identifying women at increased risk of breast cancer, and take into account the number of years of follow up each patient has within the given time period being analysed. They are summarised in *Table 29* (section 4.2.10.) and shown again in *Table 32* (below) for reference. They demonstrate that overall, the guidance differentiates low risk women from those who are at some increased risk of breast cancer. However, the difference in breast cancer rates between moderate and high risk women is not significant at all, suggesting poor differentiation (although individual analysis would suggest that high risk women are at a slightly greater risk than moderate risk women). What we can see however, is that the risk for women in either the moderate or high risk groups, excluding *BRCA* carriers, is significant from the age of 50 onwards and not significant prior to this. Even on analysis of the high risk group, the rate difference was only significant from 40-49 years when *BRCA* carriers were included in analysis, and was lost when they were excluded. This would indicate that *BRCA* carriers account for a major part of the breast cancer risk prior to age 50 years. Given that screening is offered to all moderate and high risk women from age 40 onwards, this is an important finding. In addition, it is important to note that the genuine risk increase for the moderate group occurs after the age of 50. At this age, they can be *considered* for annual screening instead of the NBSP, rather than being offered it as a standard recommendation.

Table 32. KM survival analysis data summary

NICE risk groups being compared	KM Log-Rank (*p*-value)			
	Period of follow up			
	Total follow up time	Age ≤39 years	Age 40-49 years	Age 50-59 years
Low & moderate	**0.048**	0.341	0.431	**0.037**
Low & high	**0.003**	0.091	**0.036**	0.149
Low & high (*BRCA* carriers excluded)	**0.019**	0.085	0.136	0.145
Moderate & high	0.274	0.328	0.183	0.581
Moderate & high (*BRCA* carriers excluded)	0.644	0.299	0.499	0.598
Low & moderate/high	**0.011**	0.216	0.134	**0.050**
Low & moderate/high (*BRCA* carriers excluded)	**0.024**	0.217	0.241	**0.049**

5.10. NICE guidelines overestimate the risk of breast cancer below the age of 50 years

In the meta-analysis by Pharoah et al. absolute risks were not determined, however, looking at RR, there were mixed results amongst studies regarding the risk to young women dependant on their FH (149). Some studies in the meta-analysis indicated a greater risk to young women when a relative was affected at a younger age, however whether or not this risk would meet an acceptable level to justify screening in the NHS is unclear. The results of this research suggest the risk of breast cancer below age 50, according to NICE risk category, appears to be less than suggested. The % 10-year absolute risk did not meet screening threshold aged <50 years, suggesting that there is likely to be unnecessary screening occurring in these women.

The only group in whom there was evidence of a significantly increased risk of breast cancer prior to age 50 were the *BRCA* mutation carriers. This is unsurprising given the well-established, greatly increased risk of breast cancer posed to them. Early screening in *BRCA* carriers would appear to be justified. In addition, the absolute risk to *BRCA* carriers after the age of 50 appears to increase in this cohort, contrary to findings from other studies (190). However, it is worth noting that the total number of years follow up available for known *BRCA* carriers after the age of 50 was only 19 years, compared to 75 years follow up between ages 40-49. Therefore, it cannot be ruled out that this spurious finding.

Conversely, there is no evidence emerging from this work to suggest that screening under the age of 50 in non-mutation carriers is beneficial. Below age 50, there is no significant difference between the low risk group and the moderate, high or combined risk group. Only after the age of 50 does a risk difference emerge. Therefore, there would appear to be limited benefit in screening these women according to currently suggested guidelines.

The risk to *BRCA* carriers was significant despite there being a considerable proportion of women meeting the testing threshold that were not tested. There were also women who came from families with *BRCA* mutations who had not been tested, who had to be regarded as high risk nonetheless. Genetic testing is a patient's choice,

and therefore this cannot be avoided either in a study such as this, or in the clinical setting. Nevertheless, had more *BRCA* carriers been identified in the cohort, it would still be expected that their risk prior to age 50 was significant. If more women had been tested and identified as having a mutation, they would have been analysed separately from the high risk group. The anticipated results of this would be that the overall risk in the high risk group would decrease on removal of more *BRCA* carriers. Therefore, the risks to NICE high risk, non-carrier women calculated in this study would be even less significant. Although this is theoretical, it is logical to assume that the results would remain unaffected.

The criteria appear to significantly overestimate the risk to women under the age of 50, and in addition, leave room for varied interpretation. It is unclear how the criteria corresponding to the absolute risks were determined by NICE, and the evidence presented here suggests that it should be reviewed. The fact that there is room for interpretation in the guidance is an issue in itself as it will inevitably result in a form of 'postcode lottery' for patients, in which their treatment may vary depending on which genetics department they are seen in. The overestimation may result in unnecessary screening of a large number of women, whose absolute risk of breast cancer is not equal to that of the suggested screening threshold. The discrepancies found in this study are significant, with the 95% CI of the absolute 10-year risks failing to cross the risk thresholds set by NICE. NHS standards state that for a screening programme to be suitable, the benefits gained by individuals should outweigh harms such as overtreatment, and that it should be economically balanced (229). These needs should be met to ensure appropriate resource division of NHS funding. In this case, with no significant risk above that of the low risk group identified in non-*BRCA* carriers aged less than 50, both of these requirements are unlikely to be met. The benefit of putting these patients through annual mammograms from age 40 should be questioned, especially considering both the radiation exposure and the likely anxiety induced by attending screening. The criteria being used to categorise women, and the cost-effectiveness of the current screening programme should be questioned.

5.11. Implications for current family history screening guidelines

The results presented here would imply that the NICE FH criteria used to assess breast cancer risk should be revised. It was interesting to note that in this cohort, average age of breast cancer diagnosis amongst family members did not seem to affect breast cancer risk despite being a discriminator used in the NICE guidance. At present, the current system of categorisation doesn't identify individuals at significantly increased risk over population with reliable specificity. No difference could be found in breast cancer rates between the moderate and high risk women excluding *BRCA* carriers, however when the groups are combined, the breast cancer rate is significantly greater than the low risk group between the ages of 50-59 years. Therefore, perhaps it would be beneficial to consider women without a *BRCA* mutation who meet moderate or high risk criteria as one group. Increased screening for these women would seem to be of benefit after the age of 50, and could perhaps remain more frequent than the 3-yearly screening offered as part of the NBSP. *BRCA* carriers on the other hand seem to both be identifiable using the MMSS and benefit from screening at a younger age.

Of course, more work with an adequate cohort size would be required to generate substantial evidence before the consideration of any kind of policy change. However, it is possible that the kind of approach outlined above would be considerably more cost-effective than current practise, should the results of this pilot study be replicated. Better yet, if the guidance could be modified to improve specificity at <50 years, truly high risk women could be identified. Inevitably, screening programmes are never perfect and some young onset breast cancers will be missed for a variety of reasons. However, it is important to consider the cost to society of funding a screening programme with a PPV of just 2.88 % (95% CI, 1.93-4.25%).

5.12. The family history screening programme picks up a small percentage of young onset breast cancers in Tayside

After identifying that there were 1,075 cases of female breast cancer diagnosed at <50 years in Tayside, it was determined that just 52 (4.34%) of these patient had been assessed in clinical genetics prior to their cancer diagnosis. In addition, the mean ages between the two groups of patients were very similar (42.46 for the assessed group and 42.85 for the unassessed group), suggesting that the women not seen by clinical genetics weren't developing breast cancer significantly later than those who were. Since the aim of the screening programme is to identify women who will develop young onset breast cancer, this perhaps seems like a small pick-up rate, and is an interesting observation. However, the reason for this is not clear from these results, and could be due to a number of factors. Although it is likely that breast cancers occurring at a very young age have a significant genetic component, this study did not determine whether or not the women who were not assessed in genetics prior to diagnosis had a FH of breast cancer. Even if they did, they may not have been aware of the potential risk, never made their general practitioner aware or may have simply not wished to be referred to genetics. Alternatively, if there was no FH, it may be that there were strong hormonal, environmental and lifestyle risk factors predisposing these women to disease.

These results leave a very open question of why so many young onset breast cancers are not being picked up by screening; are there simply no identifiable factors present which could indicate increased risk, or are these women simply not seeking, or being offered referral concerning their risk? In addition, at present there has been no assessment made of possible survival benefit of having been seen by a clinical geneticist. Having identified these two groups of patients, it would be useful to see whether or not patients who receive screening gain a survival benefit as a result of earlier diagnosis.

5.13. A larger cohort may be necessary to accurately assess absolute risk of breast cancer for women in each NICE risk group

Retrospective sample size calculation confirmed that this study was marginally underpowered to compare the low vs. high and moderate vs. high risk group, and significantly underpowered to compare the low vs. moderate risk groups. It is necessary to factor this in when interpreting results. To detect a clinical difference, 4000 women at either low or moderate risk of breast cancer, with clinical follow up between the ages of 40-49 would be required, which is unlikely to have been feasible in Tayside in the time frame for this research. This analysis will help inform any conclusions that are to be drawn from the study's results.

5.14. Strengths and limitations of this study

5.14.1. Limitations

This study was performed retrospectively, which is not the ideal way of performing such a cohort study. Women were selected for having no personal history of cancer, and the FH information used was that which they reported at their first appointment with clinical genetics. There was no indication at the time of FH gathering for the study which patients went on to develop cancer. Therefore, the risk of recall bias should not necessarily be a significant factor. It is possible that when there is a FH of breast cancer, the patient's anxiety may be heightened, however, over-reporting of cancer history is thought to be rare (225). The family histories used were a snapshot in time at initial assessment – changes in FH across time were not taken into account. Due to this, there may be some women for whom the risk category changed over the course of follow up. This could have been taken into account more easily using a prospective study design, which recruited patients and followed them up at given intervals. However, it is also important to acknowledge that the initial risk assessment will have been made with the relevant FH at the time, and this study assesses how the guidelines perform based on that assessment.

The fact that there was not always adequate information from clinical notes, somewhat limited the potential sample size of the study. If the study had been designed in a prospective way this could have been largely avoided by specific FH protocols. However, this may not be an accurate reflection of what happens in clinic. Time limitations meant it would not have been possible to accumulate a period of follow up time prospectively which was substantial enough to address the study aims.

It was necessary to interpret the NICE guidelines in a literal sense, meaning that clinical judgement which would most likely impact the risk categorisation of women was lost. As a result, some women may have been assigned risk categories which they otherwise would not have been in. However, it was necessary to use a standardised approach, and this flaw somewhat reflects some of the previously discussed interpretation issues which are present in the guidelines. In addition, it was

unfortunate that there was no in depth analysis of the *in situ* carcinoma development in the study for reasons outlined in section 3.5. In the UK, *in situ* carcinoma is usually treated and therefore it would not be possible to assess which would develop into invasive cancers. To do this would be considered unethical. Nevertheless, the screening programme aims to prevent deaths from invasive cancer, therefore *in situ* cancer rates are not necessarily the primary research question.

Breast cancer rates were analysed using three broad groups - <39 years, 40-49 years and 50-59 years, and no breast cancers developed over the age of 60. This means that the risks calculated reflect a rather broad age range. There are likely to be differences in risk across 5-year age brackets such as 40-44 and 45-49, which may have been beneficial to assess. However, it was necessary to consider the sample size and the number of years follow up available between each age bracket. 10-year risk assessment was necessary to provide a more statistically viable analysis. With a larger cohort size, it may be possible to look at narrower age range.

Probably the biggest limitation of this study was that it was found to be significantly underpowered. The results therefore must be interpreted with caution. In addition, the power calculation was performed using population risk figures, as the low risk group are said to be at 'population risk'. However, it is very possible that women in the low risk group are at an inherently higher risk of breast cancer nonetheless, since having any relative affected with breast cancer has been shown to increase risk by up to 2-fold, though not necessarily at a younger age (149).This would mean that the difference being detected between the low and moderate risk group, is in fact smaller than population and moderate risk. Therefore, a larger sample size would be required to detect this difference if a cohort who attended clinical genetics was used, as in this study. Regardless, as a pilot study it demonstrates methodology which could easily be applied to a larger cohort in order to produce results which can be relied upon more confidently.

5.14.2 Strengths

Since the study was looking to assess a small % difference in absolute risk of breast cancer, particularly between the low and moderate risk women, it was inevitably underpowered, as discussed above. Regardless, it does have a total patient follow up time of 15, 414 years, a considerable amount. Whilst it would be unscientific to draw definitive conclusions, it can be reasonably argued that the sample size is large enough to raise serious questions about the effectiveness of the guidelines which warrant further investigation. In addition, the 95% CIs determined when calculating the absolute % 10-year risk of breast cancer between ages 40-49, did not cross the risk thresholds for screening set by NICE.

Detailed FH information was gathered for each individual. This has built a database of FH structures which may prove beneficial in future research. There were a significant number of women in each NICE risk group which allowed comparisons to made between them. In addition, the groups were assessed to see if there were significant differences in the age at presentation which may bias the results, which there were not. The analysis also allowed fair comparison of breast cancer rates by looking at cancers per woman year of follow of time.

Overall, the methodology presented in this study provided a reasonable way of analysing the available data and assessing the effectiveness of the guidelines. Without the resource of an extended time period, the approach taken balanced the best study design and feasibility as adequately as possible. Much of the analysis could have benefitted from a larger cohort size, but this research has demonstrated that it possible to address the study aims through relatively simple means.

5.15. Conclusions

The risk of breast cancer under the age of 50 seems to be significantly greater for those with a *BRCA* mutation but, crucially, not for other moderate or high risk women in the cohort compared to the low risk group. There was no evidence to suggest a significantly increased risk of breast cancer in the high risk, non-*BRCA* mutation carriers compared to moderate risk women.

The results suggest that the combined moderate/high risk group are at an increased risk compared to the low risk group, albeit that this risk emerges after the age of 50 years, when all women would have already entered into the NBSP. Evidence to suggest a benefit of using NICE risk criteria to identify women who should be screened prior to age 50, appears to be lacking.

The current screening picks up cancers in the screened population with good sensitivity, however only a small proportion of the total young onset breast cancer in the general population is identified, and the specificity of screening is poor.

This study demonstrates simple methodology which could feasibly be expanded to other treatment centres in order to generate a sample size large enough to definitively address the research question. If replicated, the results presented here could have implications for breast cancer screening and how at-risk women are identified.

5.16. Future work

5.16.1 Increased sample size

The first thing to be addressed from the current study is the sample size. It raises a number of interesting questions about the effectiveness of the NICE guidelines, therefore it is important that they are investigated with sufficient power. It is likely that a sample size of 5000-6000 patient followed up between 40-49 years of age would be adequate. Tayside is one of the smaller health boards in Scotland, therefore recruitment of a few larger centres would most likely reach a substantial enough cohort size. The information gathering stage of this study took approximately 7 months part-time, so it would be feasible to repeat the same methodology in larger centres in a reasonably short time frame. The gold standard for a cohort study such as this would be for it to be performed prospectively. The same number of patients would be required, but to gather the required amount of follow-up time, spanning the required age ranges, would take several years. The FH reporting may be recorded more accurately, but the time frame is significantly longer and limits the speed at which results which may inform clinical practise are generated. If this study design was expanded, and the results replicated, a prospective study would be justified to generate a large, reliable database to identify family histories which truly give a significant increase in early onset breast cancer risk. To repeat the study would require multiple centres, and communication with specialists to come to a standardised way of interpreting NICE guidance.

5.16.2. Survival and cost-benefit analysis

A very important aspect missing from this study is survival analysis. The point of the screening programme is to reduce mortality in women with an increased cancer risk, rather than just identify them. Within the original FH cohort, there were simply not enough cancer diagnoses to analyse survival. However, with the addition of the cancers that were identified in women out-with the FH cohort, or who were seen prior to the year 2000, there may be a substantial enough number to carry this out. It would be very beneficial to analyse survival outcomes of women who develop breast cancer at less than 50 years, between screened and unscreened populations. Records of tumour stage at diagnosis could be analysed to see if screened cancers

were identified at an earlier stage. The most important analysis however is likely to be disease-free and overall survival from breast cancer. Given the large number of patients identified in Tayside with breast cancer <50 years, this wasn't possible within the time frame of this study. However, since the patients have been identified, it is well within the realms of feasibility for this to be carried out.

If survival analysis can be performed, it should also by extension be possible to perform a cost-benefit analysis of the screening programme. Any screening programme needs to have a trade-off between the cost to run it and the overall benefit to patients. Preliminary results presented here have suggested that the specificity of the screening programme may need to be improved. However, if the survival benefit to patients was significant enough, despite its cost, it would give more weight to the overall viability of the screening programme. Therefore, cost-benefit analysis is another very important aspect of assessing the NICE guidelines which has been omitted from this study.

5.16.3. Improving the specificity of the guidance

The specificity of the guidance in this cohort was poor, therefore it would be interesting to do work which aimed to improve this. Given that detailed FH information was recorded for each patient, it may be possible to assess the exact breast cancer risks associated with differing family histories. This may enable remodelling using similar empiric FH structures as NICE, but with improved specificity. It would be desirable to maintain a good level of sensitivity, but the overall cost-effectiveness of the screening programme is likely to be improved if specificity is increased. In addition, being able to do this relying on simple FH parameters such as number and age of relatives affected would be more user-friendly for clinicians than a model which requires specific pathology or lifestyle information, since this isn't always available. This study collected the dataset that would be required to do this, so by extending it to other centres and increasing its size, a large enough sample to investigate alternative criteria would be available.

6.0. References

1. Cancer Research UK. Breast cancer statistics. [cited 2016 November]; Available from: http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/breast-cancer#bxHjcylJv1CGLRrR.99
2. Cancer Research UK. Breast cancer 5-year survival. [cited 2016 November]; Available from: http://www.cancerresearchuk.org/sites/default/files/cstream-node/surv_5yr_age_breast_0.pdf
3. NHS Screening Programmes. Breast screening age extension trial for women who are younger than 50 or older than 70. In: Public Health England, editor.; 2016.
4. Marmot MG, Altman DG, Cameron DA, Dewar JA, Thompson SG, Wilcox M. The benefits and harms of breast cancer screening: an independent review. Lancet. 2012 October 2012;380(9855):1778-86.
5. Pharoah PD. Cost effectiveness of the NHS breast screening programme: life table model. BMJ. 2013;346(f2618).
6. Li Cl, Anderson BO, Daling JR, Moe RE. Trends in incidence rates of invasive lobular and ductal breast carcinoma. JAMA. 2003 March;289(11):1421-4.
7. Dossus L, Benusiglio PR. Lobular breast cancer: incidence and genetic and non-genetic risk factors. Breast Cancer Res. 2015;17(37).
8. Wu VS, Kanaya N, Lo C, Mortimer J, Chen S. From bench to bedside: What do we know about hormone receptor-positive and human epidermal growth factor receptor 2-positive breast cancer? J Steroid Biochem Mol Biol. 2015;153:45-53.
9. Banerjee M, George J, Song EY, Roy A, Hryniuk W. Tree-based model for breast cancer prognostication. J Clin Oncol. 2004 Jul 01;22(13):2567-75.
10. Purdie CA, Quinlan P, Jordan LB, Ashfield A, Ogston S, Dewar JA, et al. Progesterone receptor expression is an independent prognostic variable in early breast cancer: a population-based study. Br J Cancer. 2014 Feb 04;110(3):565-72.
11. Anderson WF, Chatterjee N, Ershler WB, Brawley OW. Estrogen receptor breast cancer phenotypes in the Surveillance, Epidemiology, and End Results database. Breast Cancer Res Treat. 2002 November;76(1):27-36.
12. Anderson WF, Rosenberg PS, Prat A, Perou CM, Sherman ME. How many etiological subtypes of breast cancer: two, three, four, or more? J Natl Cancer Inst. 2014 August;106(8).
13. Arteaga CL, Sliwkowski MX, Osborne CK, Perez EA, Puglisi F, Gianni L. Treatment of HER2-positive breast cancer: current status and future perspectives. Nat Rev Clin Oncol. 2011 November;9(1):16-32.
14. Evans DG, Lalloo F, Howell S, Verhoef S, Woodward ER, Howell A. Low prevalence of HER2 positiviyu amongst *BRCA1* and *BRCA2* mutation carriers and in primary *BRCA* screens. Breast Cancer Res Treat. 2016 February;155(3):597-601.
15. Kumar P, Aggarwal R. An overview of triple-negative breast cancer. Arch Gynecol Obstet. 2016 February;293(2):247-69.
16. Cadoo KA, Fornier MN, Morris PG. Biological subtypes of breast cancerL current concepts and implications for recurrence patterns. Q J Nucl Med Mol Imaging. 2013 December;57(4):312-21.
17. The Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. Nature. 2012 Oct 04;490(7418):61-70.
18. Barnard ME, Boeke CE, Tamimi RM. Established breast cancer risk factors and risk of intrinsic subtypes. Biochim Biophys Acta. 2015 August;1856(1):73-85.

19. Dai X, Li T, Bai Z, Yang Y, Liu X, Zhan J, et al. Breast cancer intrinsic subtype classification, clinical use and future trends. Am J Cancer Res. 2015 September;5(10):2929-43.
20. Klassen AC, Smith KC. The enduring and evolving relationship between social class and breast cancer burden: a review of the literature. Cancer Epidemiol. 2011 Jun;35(3):217-34.
21. Carnon AG, Ssemwogerere A, Lamont DW, Hole DJ, Mallon EA, George WD, et al. Relation between socioeconomic deprivation and pathological prognostic factors in women with breast cancer. BMJ. 1994 Oct 22;309(6961):1054-7.
22. Orsini M, Tretarre B, Daures JP, Bessaoud F. Individual socioeconomic status and breast cancer diagnostic stages: a French case-control study. Eur J Public Health. 2016 Jun;26(3):445-50.
23. Mohaghegh P, Yavari P, Akbari ME, Abadi A, Ahmadi F. Associations of demographic and socioeconomic factors with stage at diagnosis of breast cancer. Asian Pac J Cancer Prev. 2015;16(4):1627-31.
24. Reeder-Hayes KE, Wheeler SB, Mayer DK. Health disparities across the breast cancer continuum. Semin Oncol Nurs. 2015 May;31(2):170-7.
25. Kohler BA, Sherman RL, Howlader N, Jemal A, Ryerson AB, Henry KA, et al. Annual Report to the Nation on the Status of Cancer, 1975-2011, Featuring Incidence of Breast Cancer Subtypes by Race/Ethnicity, Poverty, and State. J Natl Cancer Inst. 2015 Jun;107(6):djv048.
26. Dunn BK, Agurs-Collins T, Browne D, Lubet R, Johnson KA. Health disparities in breast cancer: biology meets socioeconomic status. Breast Cancer Res Treat. 2010 Jun;121(2):281-92.
27. DeSantis C, Ma J, Bryan L, Jemal A. Breast cancer statistics, 2013. CA Cancer J Clin. 2014 Jan-Feb;64(1):52-62.
28. Miller JW, King JB, Joseph DA, Richardson LC. Breast cancer screening among adult women--Behavioral Risk Factor Surveillance System, United States, 2010. MMWR Suppl. 2012 Jun 15;61(2):46-50.
29. Bowen RL, Stebbing J, Jones LJ. A review of the ethnic differences in breast cancer. Pharmacogenomics. 2006 Sep;7(6):935-42.
30. Singh RR, Kumar R. Steroid hormone receptor signaling in tumorigenesis. J Cell Biochem. 2005 October;96(3):490-505.
31. Trivers KF, Lund MJ, Porter PL, Liff JM, Flagg EW, Coates RJ, et al. The epidemiology of triple-negative breast cancer, including race. . Cancer Causes Control. 2009 September;20(7):1071-82.
32. Abdulkareem IH. Aetio-pathogenesis of breast cancer. Niger Med J. 2013 November;54(6):371-5.
33. Dumitrescu RG, Cotarla I. Understanding breast cancer risk -- where do we stand in 2005? J Cell Mol Med. 2005 March;9(1):208-21.
34. Hsieh CC, Trichopoulos D, Katsouyanni K, S. Y. Age at menarche, age at menopause, height and obesity as risk factors for breast cancerL associations and interactions in an international case-control study. . Int J Cancer. 1990 November;46(5):796-800.
35. Britt K, Ashworth A, Smalley M. Pregnancy and the risk of breast cancer. Endocr Relat Cancer. 2007 December;14(4):907-33.
36. Malara NM, Leotta A, Sidoti A, Lio S, D'Angelo R, Caparello B, et al. Ageing, hormonal behaviour and cyclin D1 in Ductal breast carcinomas. Breast 2006 February;15(1):81-9.
37. Million Women Study Collaborators. Breast cancer and hormone-replacement therapy in the Million Women Study. Lancet. 2003 August;362(9382):419-27.
38. National Institute for Health and Care Excellence. Menopause: diagnosis and management NG23. NICE; 2015.

39. Gierisch JM, Coeytaux RR, Urratia RP, Havrilesky LJ, Moorman PG, Lowery WJ, et al. Oral contraceptive use and risk of breast, cervical, colorectal, and endometrial cancers: a systematic review. Cancer Epidemiol Biomarkers Prev. 2013 November;22(11):1931-43.
40. National Institute for Health and Care Excellence. Contraception - combined hormonal methods. NICE; 2016.
41. Moorman PG, Havrilesky LJ, Gierisch JM, Coeytaux RR, Lowery WJ, Peragallo-Urrutia R, et al. Oral contraceptives and risk of ovarian cancer and breast cancer among high-risk women: a systematic review and meta-analysis J Clin Oncol. 2013 November;31(33):188-98.
42. Anderson KN, Schwab RB, Martinez ME. Reproductive risk factors and breast cancer subtypes: a review of the literature. Breast Cancer Res Treat. 2014 February;144(1):1-10.
43. Sherman ME, Rimm DL, Yang XR, Chatterjee N, Brinton LA, Lissowska J, et al. Variation in breast cancer hormone receptor and HER2 levels by etiologic factors: a population-based analysis. Int J Cancer. 2007 September;121(5):1079-85.
44. Trentham-Dietz A, Newcomb PA, Nichols HB, Hampton JM. Breast cancer risk factors and second primary malignancies among women with breast cancer. Breast Cancer Res Treat. 2007 October;105(2):195-207.
45. Huang Z, Hankinson SE, Colditz GA, Stampfer MJ, Hunter DJ, Manson JE, et al. Dual effects of weight and weight gain on breast cancer risk. JAMA. 1997 November;278(17):1407-11.
46. Wang J, Yang DL, Chen ZZ, Gou BF. Associations of body mass index with cancer incidence among populations, genders, and menopausal status: A systematic review and meta-analysis. Cancer Epidemiol. 2016 June;42(1):1-8.
47. Bulun SE, Simpson ER. Competitive reverse transcription-polymerase chain reaction analysis indicates that levels of aromatase cytochrome P450 transcripts in adipose tissue of buttocks, thighs, and abdomen of women increase with advancing age. . JCEM. 2009;78(2).
48. Goodwin PJ, Ambrosone CB, Hong CC. Modifiable Lifestyle Factors and Breast Cancer Outcomes: Current Controversies and Research Recommendations. Adv Exp Med Biol. 2015;862(1):177-92.
49. Harvie M, Howell A, Evans DG. Can diet and lifestyle prevent breast cancer: what is the evidence? In: DS D, editor. American Society of Clinical Oncology Educational Book. Chicago, Illinois: ASCO; 2015. p. e66-73.
50. Aune D, Chan DS, Greenwood DC, Vieira AR, Rosenblatt DA, Vieira R, et al. Dietary fiber and breast cancer risk: a systematic review and meta-analysis of prospective studies. Ann Oncol. 2012 June;23(5):1394-402.
51. Chlebowski RT, Rose D, Buzzard IM, Blackburn GL, Insull W Jr, Grosvenor M, et al. Adjuvant dietary fat intake reduction in postmenopausal breast cancer patient management. The Women's Intervention Nutrition Study (WINS). Breast Cancer Res Treat. 1992 January;20(2):73-84.
52. Pierce JP, Natarajan L, Caan BJ, Parker BA, Greenberg ER, Flatt SW, et al. Influence of a diet very high in vegetables, fruit, and fiber and low in fat on prognosis following treatment for breast cancer: the Women's Healthy Eating and Living (WHEL) randomized trial. JAMA. 2007 July;298(2):289-98.
53. Dong JY, Qin LQ. Soy isoflavones consumption and risk of breast cancer incidence or recurrence: a meta-analysis of prospective studies. Breast Cancer Res Treat. 2011 January;125(2):315-23.
54. Lynch BM, Neilson HK, Friedenreich CM. Physical activity and breast cancer prevention. Recent Results Cancer Res. 2011;186(1):13-42.
55. Liu Y, Nguyen N, Colditz GA. Links between alcohol consumption and breast cancer: a look at the evidence. Womens Health (Lond). 2015 January;11(1):65-77.

56. Reichman ME, Judd JT, Longcope C, Schatzkin A, Clevidence BA, Nair PP, et al. Effects of alcohol consumption on plasma and urinary hormone concentrations in premenopausal women. J Natl Cancer Inst. 1993 May;85(9):722-7.
57. Onland-Moret NC, Peeters PH, van der Schouw YT, Grobbee DE, van Gils CH. Alcohol and endogenous sex steroid levels in postmenopausal women: a cross-sectional study. J Clin Endocrinol Metab. 2005 March;90(3):1414-9.
58. Singletary KW, Frey RS, Yan W. Effect of ethanol on proliferation and estrogen receptor-alpha expression in human breast cancer cells. Cancer Lett. 2001 April;165(2):131-7.
59. Gandini S, Botteri E, Iodice S, Boniol M, Lowenfels AB, Maisonneuve P, et al. Tobacco smoking and cancer: a meta-analysis. Int J Cancer. 2008 January;122(1):155-64.
60. Reynolds P. Smoking and breast cancer. J Mammary Gland Biol Neoplasia 2013 March;18(1):15-23.
61. Marcus PM, Newman B, Millikan RC, Moorman PG, Baird DD, Qaqish B. The associations of adolescent cigarette smoking, alcoholic beverage consumption, environmental tobacco smoke, and ionizing radiation with subsequent breast cancer risk. Cancer Causes Control. 2000 March;11(3):271-78.
62. Pettersson A et al. Mammographic density phenotypes and risk of breast cancer: a meta-analysis. J Natl Cancer Inst. 2014 May;106(5).
63. McCormack VA, dos Santos Silva I. Breast density and parenchymal patterns as markers of breast cancer risk: a meta-analysis. Cancer Epidemiol Biomarkers Prev. 2006 June;15(6):1159-69.
64. McInnes KJ, Brown KA, Knower KC, Chand AL, Clyne CD, Simpson ER. Characterisation of aromatase expression in the human adipocyte cell line SGBS. Breast Cancer Res Treat. 2008 December;112(3):429-35.
65. Brower V. Homing in on mechanisms linking breast density to breast cancer risk. J Natl Cancer Inst. 2010 June;102(12):843-5.
66. Howell A, Anderson AS, Clarke RB, Duffy SW, Evans DG, Garcia-Closas M, et al. Risk determination and prevention of breast cancer. Breast Cancer Res. 2014 September;16(5).
67. Huo CW, Chew GL, Britt KL, Ingman WV, Henderson MA, Hopper JL, et al. Mammographic density-a review on the current understanding of its association with breast cancer. Breast Cancer Res Treat. 2014 April;144(3):479-502.
68. Colin C, Schott AM, Valette PJ. Mammographic density is not a worthwhile examination to distinguish high cancer risk women in screening. Eur Radiol. 2014 October;24(10):2412-6.
69. 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 September;98(17):1215-26.
70. Tice JA, Cummings SR, Ziv E, Kerlikowske K. Mammographic breast density and the Gail model for breast cancer risk prediction in a screening population. Breast Cancer Res Treat. 2005 November;94(2):115-22.
71. Drooger J, Hooning MJ, Seynaeve CM, Baaijens MH, Obdeijn IM, Sleijfer S, et al. Diagnostic and therapeutic ionizing radiation and the risk of a first and second primary breast cancer, with special attention for *BRCA1* and *BRCA2* mutation carriers: a critical review of the literature. Cancer Treat Rev. 2015 February;41(2):187-96.
72. Early Breast Cancer Trialists' Collaborative Group (EBCTCG). Effect of radiotherapy after breast-conserving surgery on 10-year recurrence and 15-year breast cancer death: meta-analysis of individual patient data for 10 801 women in 17 randomised trials. Lancet. 2011 November;378(9804):1707-16.
73. Preston DL, Ron E, Tokuoka S, Funamoto S, Nishi N, Soda M, et al. Solid cancer incidence in atomic bomb survivors: 1958-1998. Radiat Res. 2007 July;168(1):1-64.

74. Swerdlow AJ, Cooke R, Bates A, Cunningham D, Falk SJ, Gilson D, et al. Breast cancer risk after supradiaphragmatic radiotherapy for Hodgkin's lymphoma in England and Wales: a National Cohort Study. J Clin Oncol. 2012 August;30(22):2745-52.
75. Hoffman DA, Lonstein JE, Morin MM, Visscher W, Harris BS 3rd, Boice JD Jr. Breast cancer in women with scoliosis exposed to multiple diagnostic x rays. J Natl Cancer Inst. 1989 September;81(17):1307-12.
76. Eidemüller M, Holmberg E, Jacob P, Lundell M, Karlsson P. Breast cancer risk and possible mechanisms of radiation-induced genomic instability in the Swedish hemangioma cohort after reanalyzed dosimetry. Mutat Res. 2015 May;775(1):1-9.
77. Gao X, Fisher SG, Emami B. Risk of second primary cancer in the contralateral breast in women treated for early-stage breast cancer: a population-based study. Int J Radiat Oncol Biol Phys. 2003 July;56(4):1038-45.
78. Darby S, McGale P, Correa C, Taylor C, Arriagada R, Clarke M, et al. Effect of radiotherapy after breast-conserving surgery on 10-year recurrence and 15-year breast cancer death: meta-analysis of individual patient data for 10,801 women in 17 randomised trials. Lancet. 2011 Nov 12;378(9804):1707-16.
79. National Institute for Health and Care Excellence. Familial breast cancer: classification, care and managing breast cancer and related risks in people with a family history of breast cancer CG164. NICE; 2013.
80. Pijpe A, Andrieu N, Easton DF, Kesminiene A, Cardis E, Noguès C, et al. Exposure to diagnostic radiation and risk of breast cancer among carriers of *BRCA1/2* mutations: retrospective cohort study (GENE-RAD-RISK). BMJ. 2012 September;345(e5660).
81. U.S. Preventative Services Task Force. Genetic risk assessment and *BRCA* mutation testing for breast and ovarian cancer susceptibility: recommendation statement. Ann Intern Med. 2005 Sep 6;143(5):355-61.
82. Meindl A, Ditsch N, Kast K, Rhiem K, Schmutzler RK. Hereditary breast and ovarian cancer: new genes, new treatments, new concepts. Dtsch Arztebl Int. 2011 May;108(19):323-30.
83. Turnbull C, Rahman N. Genetic predisposition to breast cancer: past, present, and future. Annu Rev Genomics Hum Genet. 2008;9:321-45.
84. Pharoah PD, Stratton JF, Mackay J. Screening for breast and ovarian cancer: the relevance of family history. Br Med Bull. 1998;54(4):823-38.
85. Hall JM, Lee MK, Newman B, Morrow JE, Anderson LA, Huey B, et al. Linkage of early-onset familial breast cancer to chromosome 17q21. Science. 1990 Dec 21;250(4988):1684-9.
86. 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.
87. 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.
88. Ma JL, Kim EM, Haber JE, Lee SE. Yeast Mre11 and Rad1 proteins define a Ku-independent mechanism to repair double-strand breaks lacking overlapping end sequences. Mol Cell Biol. 2003 Dec;23(23):8820-8.
89. Bhattacharyya A, Ear US, Koller BH, Weichselbaum RR, Bishop DK. The breast cancer susceptibility gene *BRCA1* is required for subnuclear assembly of Rad51 and survival following treatment with the DNA cross-linking agent cisplatin. J Biol Chem. 2000 Aug 4;275(31):23899-903.
90. Moynahan ME, Pierce AJ, Jasin M. *BRCA2* is required for homology-directed repair of chromosomal breaks. Mol Cell. 2001 Feb;7(2):263-72.
91. Gudmundsdottir K, Ashworth A. The roles of *BRCA1* and *BRCA2* and associated proteins in the maintenance of genomic stability. Oncogene. 2006 Sep 25;25(43):5864-74.

92. Bertwistle D, Ashworth A. The pathology of familial breast cancer: How do the functions of *BRCA1* and *BRCA2* relate to breast tumour pathology? Breast Cancer Research: BCR. 1999;1(1):41-7.
93. Anglican Breast Cancer Study Group. Prevalence and penetrance of *BRCA1* and *BRCA2* mutations in a population-based series of breast cancer cases. . British Journal Of Cancer. 2000 Nov;83(10):1301-8.
94. 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. American Journal Of Human Genetics. 2003 May;72(5):1117-30.
95. Lakhani SR, Van De Vijver MJ, Jacquemier J, Anderson TJ, Osin PP, McGuffog L, et al. The pathology of familial breast cancer: predictive value of immunohistochemical markers estrogen receptor, progesterone receptor, HER-2, and p53 in patients with mutations in *BRCA1* and *BRCA2.* Journal Of Clinical Oncology: Official Journal Of The American Society Of Clinical Oncology. 2002 May 1;20(9):2310-8.
96. Inoue K, Fry EA, Frazier DP. Transcription factors that interact with p53 and Mdm2. International Journal Of Cancer Journal International Du Cancer. 2016 Apr 1;138(7):1577-85.
97. Bieging KT, Mello SS, Attardi LD. Unravelling mechanisms of p53-mediated tumour suppression. Nat Rev Cancer. 2014 May;14(5):359-70.
98. Birch JM, Alston RD, McNally RJ, Evans DG, Kelsey AM, Harris M, et al. Relative frequency and morphology of cancers in carriers of germline TP53 mutations. Oncogene. 2001 Aug 2;20(34):4621-8.
99. Wilson JR, Bateman AC, Hanson H, An Q, Evans G, Rahman N, et al. A novel HER2-positive breast cancer phenotype arising from germline TP53 mutations. J Med Genet. 2010 Nov;47(11):771-4.
100. Leslie NR, Longy M. Inherited PTEN mutations and the prediction of phenotype. Semin Cell Dev Biol. 2016 Apr;52:30-8.
101. Bigner SH, Mark J, Burger PC, Mahaley MS, Jr., Bullard DE, Muhlbaier LH, et al. Specific chromosomal abnormalities in malignant human gliomas. Cancer Res. 1988 Jan 15;48(2):405-11.
102. Li J, Yen C, Liaw D, Podsypanina K, Bose S, Wang SI, et al. PTEN, a putative protein tyrosine phosphatase gene mutated in human brain, breast, and prostate cancer. Science. 1997 Mar 28;275(5308):1943-7.
103. Liaw D, Marsh DJ, Li J, Dahia PL, Wang SI, Zheng Z, et al. Germline mutations of the PTEN gene in Cowden disease, an inherited breast and thyroid cancer syndrome. Nat Genet. 1997 May;16(1):64-7.
104. Uppal S, Mistry D, Coatesworth AP. Cowden disease: a review. Int J Clin Pract. 2007 Apr;61(4):645-52.
105. Suzuki A, Nakano T, Mak TW, Sasaki T. Portrait of PTEN: messages from mutant mice. Cancer Sci. 2008 Feb;99(2):209-13.
106. Song MS, Salmena L, Pandolfi PP. The functions and regulation of the PTEN tumour suppressor. Nat Rev Mol Cell Biol. 2012 Apr 04;13(5):283-96.
107. Economopoulou P, Dimitriadis G, Psyrri A. Beyond *BRCA*: new hereditary breast cancer susceptibility genes. Cancer Treat Rev. 2015 Jan;41(1):1-8.
108. Pharoah PD, Antoniou A, Bobrow M, Zimmern RL, Easton DF, Ponder BA. Polygenic susceptibility to breast cancer and implications for prevention. Nat Genet. 2002 May;31(1):33-6.
109. Meijers-Heijboer H, van den Ouweland A, Klijn J, Wasielewski M, de Snoo A, Oldenburg R, et al. Low-penetrance susceptibility to breast cancer due to CHEK2(*)1100delC in noncarriers of *BRCA1* or *BRCA2* mutations. Nat Genet. 2002 May;31(1):55-9.

110. Guleria A, Chandna S. ATM kinase: Much more than a DNA damage responsive protein. DNA Repair (Amst). 2016 Mar;39:1-20.
111. Renwick A, Thompson D, Seal S, Kelly P, Chagtai T, Ahmed M, et al. ATM mutations that cause ataxia-telangiectasia are breast cancer susceptibility alleles. Nat Genet. 2006 Aug;38(8):873-5.
112. Peng M, Litman R, Jin Z, Fong G, Cantor SB. BACH1 is a DNA repair protein supporting *BRCA1* damage response. Oncogene. 2006 Apr 6;25(15):2245-53.
113. Levitus M, Waisfisz Q, Godthelp BC, de Vries Y, Hussain S, Wiegant WW, et al. The DNA helicase BRIP1 is defective in Fanconi anemia complementation group J. Nat Genet. 2005 Sep;37(9):934-5.
114. Seal S, Thompson D, Renwick A, Elliott A, Kelly P, Barfoot R, et al. Truncating mutations in the Fanconi anemia J gene BRIP1 are low-penetrance breast cancer susceptibility alleles. Nat Genet. 2006 Nov;38(11):1239-41.
115. Aloraifi F, McCartan D, McDevitt T, Green AJ, Bracken A, Geraghty J. Protein-truncating variants in moderate-risk breast cancer susceptibility genes: a meta-analysis of high-risk case-control screening studies. Cancer Genet. 2015 Sep;208(9):455-63.
116. Sopik V, Foulkes WD. Risky business: getting a grip on BRIP. J Med Genet. 2016 May;53(5):296-7.
117. Lavin MF, Kozlov S, Gatei M, Kijas AW. ATM-Dependent Phosphorylation of All Three Members of the MRN Complex: From Sensor to Adaptor. Biomolecules. 2015;5(4):2877-902.
118. Christou CM, Kyriacou K. *BRCA1* and Its Network of Interacting Partners. Biology (Basel). 2013;2(1):40-63.
119. Heikkinen K, Rapakko K, Karppinen SM, Erkko H, Knuutila S, Lundan T, et al. RAD50 and NBS1 are breast cancer susceptibility genes associated with genomic instability. Carcinogenesis. 2006 Aug;27(8):1593-9.
120. Tommiska J, Seal S, Renwick A, Barfoot R, Baskcomb L, Jayatilake H, et al. Evaluation of RAD50 in familial breast cancer predisposition. Int J Cancer. 2006 Jun 1;118(11):2911-6.
121. Park JY, Zhang F, Andreassen PR. PALB2: the hub of a network of tumor suppressors involved in DNA damage responses. Biochim Biophys Acta. 2014 Aug;1846(1):263-75.
122. Xia B, Dorsman JC, Ameziane N, de Vries Y, Rooimans MA, Sheng Q, et al. Fanconi anemia is associated with a defect in the *BRCA2* partner PALB2. Nat Genet. 2007 Feb;39(2):159-61.
123. Reid S, Schindler D, Hanenberg H, Barker K, Hanks S, Kalb R, et al. Biallelic mutations in PALB2 cause Fanconi anemia subtype FA-N and predispose to childhood cancer. Nat Genet. 2007 Feb;39(2):162-4.
124. Rahman N, Seal S, Thompson D, Kelly P, Renwick A, Elliott A, et al. PALB2, which encodes a BRCA2-interacting protein, is a breast cancer susceptibility gene. Nat Genet. 2007 Feb;39(2):165-7.
125. Erkko H, Xia B, Nikkila J, Schleutker J, Syrjakoski K, Mannermaa A, et al. A recurrent mutation in PALB2 in Finnish cancer families. Nature. 2007 Mar 15;446(7133):316-9.
126. Erkko H, Dowty JG, Nikkila J, Syrjakoski K, Mannermaa A, Pylkas K, et al. Penetrance analysis of the PALB2 c.1592delT founder mutation. Clin Cancer Res. 2008 Jul 15;14(14):4667-71.
127. Southey MC, Teo ZL, Dowty JG, Odefrey FA, Park DJ, Tischkowitz M, et al. A PALB2 mutation associated with high risk of breast cancer. Breast Cancer Res. 2010;12(6):R109.
128. Byrnes GB, Southey MC, Hopper JL. Are the so-called low penetrance breast cancer genes, ATM, BRIP1, PALB2 and CHEK2, high risk for women with strong family histories? Breast Cancer Res. 2008;10(3):208.

129. Buys SS, Sandbach JF, Gammon A, Patel G, Kidd J, Brown KL, et al. A study of over 35,000 women with breast cancer tested with a 25-gene panel of hereditary cancer genes. Cancer. 2017 Jan 13.
130. Maxwell KN, Nathanson KL. Common breast cancer risk variants in the post-COGS era: a comprehensive review. Breast Cancer Res. 2013 December;15(6).
131. Walsh MF, Nathanson KL, Couch FJ, Offit K. Genomic Biomarkers for Breast Cancer Risk. Adv Exp Med Biol. 2016;882(1):1-32.
132. Michailidou K et al. Large-scale genotyping identifies 41 new loci associated with breast cancer risk. Nat Genet. 2013 April;45(5):353-61.
133. National Center for Biotechnology Information. dbSNP - Short Genetic Variations: rs2981579. [cited 2016 May]; Available from: https://www.ncbi.nlm.nih.gov/SNP/snp_ref.cgi?rs=2981579
134. National Center for Biotechnology Information. dbSNP - Short Genetic Variation: rs2981582. [cited; Available from: http://www.ncbi.nlm.nih.gov/SNP/snp_ref.cgi?rs=2981582
135. Easton DF et al. Genome-wide association study identifies novel breast cancer susceptibility loci. Nature. 2007 June;447(7148):1087-93.
136. International HapMap Consortium. The International HapMap Project. Nature. 2003 December;426(6968):789-96.
137. André F, J. C. Rationale for targeting fibroblast growth factor receptor signaling in breast cancer. Breast Cancer Res Treat. 2015 February;150(1):1-8.
138. Martin AJ, Grant A, Ashfield AM, Palmer CN, Baker L, Quinlan PR, et al. FGFR2 protein expression in breast cancer: nuclear localisation and correlation with patient genotype. BMC Res Notes. 2011 March;4(1).
139. National Center for Biotechnology Information. dbSNP - Short Genetic Variations: rs10069690. [cited 2016 May]; Available from: http://www.ncbi.nlm.nih.gov/projects/SNP/snp_ref.cgi?rs=10069690
140. Long J, Zhang B, Signorello LB, Cai Q, Deming-Halverson S, Shrubsole MJ, et al. Evaluating genome-wide association study-identified breast cancer risk variants in African-American women. PLoS One. 2013 April;8(4).
141. Heidenreich B, Rachakonda PS, Hemminki K, Kumar R. TERT promoter mutations in cancer development. Curr Opin Genet Dev. 2014 February;24(1):30-7.
142. Horn S, Figl A, Rachakonda PS, Fischer C, Sucker A, Gast A, et al. TERT promoter mutations in familial and sporadic melanoma. Science. 2013 February;339(6122):959-61.
143. National Center for Biotechnology Information. dbSNP - Short Genetic Variations: rs1045485. [cited 2016 May]; Available from:
144. Cox A et al. A common coding variant in CASP8 is associated with breast cancer risk. Nat Genet. 2007 February;39(3):352-8.
145. Pharoah PD, Antoniou AC, Easton DF, Ponder BA. Polygenes, risk prediction, and targeted prevention of breast cancer. N Engl J Med. 2008 June;358(26):2796-803.
146. Gail MH, Benichou J. Validation studies on a model for breast cancer risk. J Natl Cancer Inst. 1994 Apr 20;86(8):573-5.
147. Evans DG, Brentnall A, Byers H, Harkness E, Stavrinos P, Howell A, et al. The impact of a panel of 18 SNPs on breast cancer risk in women attending a UK familial screening clinic: a case-control study. J Med Genet. 2016;doi: 10.1136/jmedgenet-2016-104125.
148. Lynch H, Wen H, Kim YC, Snyder C, Kinarsky Y, Chen PX, et al. Can unknown predisposition in familial breast cancer be family-specific? Breast J. 2013 Sep-Oct;19(5):520-8.
149. 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.

150. Collaborative Group in Hormonal Factors in Breast Cancer. Familial breast cancer: collaborative reanalysis of individual data from 52 epidemiological studies including 58,209 women with breast cancer and 101,986 women without the disease. Lancet. 2001 Oct 27;358(9291):1389-99.
151. Colditz GA, Kaphingst KA, Hankinson SE, Rosner B. Family history and risk of breast cancer: nurses' health study. Breast Cancer Res Treat. 2012 Jun;133(3):1097-104.
152. Molino A, Giovannini M, Pedersini R, Frisinghelli M, Micciolo R, Mandara M, et al. Correlations between family history and cancer characteristics in 2256 breast cancer patients. Br J Cancer. 2004 Jul 5;91(1):96-8.
153. Thalib L, Wedren S, Granath F, Adami HO, Rydh B, Magnusson C, et al. Breast cancer prognosis in relation to family history of breast and ovarian cancer. Br J Cancer. 2004 Apr 5;90(7):1378-81.
154. Duffy SW, Dibden A, Michalopoulos D, Offman J, Parmar D, Jenkins J, et al. Screen detection of ductal carcinoma in situ and subsequent incidence of invasive interval breast cancers: a retrospective population-based study. Lancet Oncol. 2016 January;17(1):109-14.
155. Vilaprinyo E, Forné C, Carles M, Sala M, Pla R, Castells X, et al. Cost-effectiveness and harm-benefit analyses of risk-based screening strategies for breast cancer. PLoS One. 2014 February;9(2).
156. 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.
157. Baker LH. Breast Cancer Detection Demonstration Project: five-year summary report. CA Cancer J Clin. 1982 Jul-Aug;32(4):194-225.
158. 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.
159. Anderson DE, Badzioch MD. Risk of familial breast cancer. Cancer. 1985 Jul 15;56(2):383-7.
160. Belanger CF, Hennekens CH, Rosner B, Speizer FE. The nurses' health study. Am J Nurs. 1978 Jun;78(6):1039-40.
161. Spiegelman D, Colditz GA, Hunter D, Hertzmark E. Validation of the Gail et al. model for predicting individual breast cancer risk. J Natl Cancer Inst. 1994 Apr 20;86(8):600-7.
162. Bondy ML, Lustbader ED, Halabi S, Ross E, Vogel VG. Validation of a breast cancer risk assessment model in women with a positive family history. J Natl Cancer Inst. 1994 Apr 20;86(8):620-5.
163. Euhus DM, Leitch AM, Huth JF, Peters GN. Limitations of the Gail model in the specialized breast cancer risk assessment clinic. Breast J. 2002 Jan-Feb;8(1):23-7.
164. Meads C, Ahmed I, Riley RD. A systematic review of breast cancer incidence risk prediction models with meta-analysis of their performance. Breast Cancer Res Treat. 2012 Apr;132(2):365-77.
165. Claus EB, Risch NJ, Thompson WD. Age at onset as an indicator of familial risk of breast cancer. Am J Epidemiol. 1990 Jun;131(6):961-72.
166. Claus EB, Risch N, Thompson WD. Genetic analysis of breast cancer in the cancer and steroid hormone study. Am J Hum Genet. 1991 Feb;48(2):232-42.
167. Claus EB, Risch N, Thompson WD. Autosomal dominant inheritance of early-onset breast cancer. Implications for risk prediction. Cancer. 1994 Feb 1;73(3):643-51.
168. Claus EB, Risch N, Thompson WD. The calculation of breast cancer risk for women with a first degree family history of ovarian cancer. Breast Cancer Res Treat. 1993 Nov;28(2):115-20.
169. Evans DG, Howell A. Breast cancer risk-assessment models. Breast Cancer Res. 2007;9(5):213.

170. Amir E, Evans DG, Shenton A, Lalloo F, Moran A, Boggis C, et al. Evaluation of breast cancer risk assessment packages in the family history evaluation and screening programme. J Med Genet. 2003 Nov;40(11):807-14.
171. Jacobi CE, de Bock GH, Siegerink B, van Asperen CJ. Differences and similarities in breast cancer risk assessment models in clinical practice: which model to choose? Breast Cancer Res Treat. 2009 May;115(2):381-90.
172. van Asperen CJ, Jonker MA, Jacobi CE, van Diemen-Homan JE, Bakker E, Breuning MH, et al. Risk estimation for healthy women from breast cancer families: new insights and new strategies. Cancer Epidemiol Biomarkers Prev. 2004 Jan;13(1):87-93.
173. Hoskins KF, Stopfer JE, Calzone KA, Merajver SD, Rebbeck TR, Garber JE, et al. Assessment and counseling for women with a family history of breast cancer. A guide for clinicians. JAMA. 1995 Feb 15;273(7):577-85.
174. 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.
175. Ford D, Easton DF, Peto J. Estimates of the gene frequency of *BRCA1* and its contribution to breast and ovarian cancer incidence. Am J Hum Genet. 1995 Dec;57(6):1457-62.
176. Andersen TI. Genetic heterogeneity in breast cancer susceptibility. Acta Oncol. 1996;35(4):407-10.
177. Shattuck-Eidens D, McClure M, Simard J, Labrie D, Narod S, Couch F, et al. A collaborative survey of 80 mutations in the *BRCA1* breast and ovarian cancer susceptibility gene. Implications for presymptomatic testing and screening. JAMA. 1995 February;273(7):535-41.
178. BayesMendel Lab. BRCAPRO. [cited 2015 November]; Available from: http://bcb.dfci.harvard.edu/bayesmendel/brcapro.php
179. Berry DA, Iversen ES, Jr., Gudbjartsson DF, Hiller EH, Garber JE, Peshkin BN, et al. BRCAPRO validation, sensitivity of genetic testing of *BRCA1/BRCA2*, and prevalence of other breast cancer susceptibility genes. J Clin Oncol. 2002 Jun 1;20(11):2701-12.
180. Euhus DM, Smith KC, Robinson L, Stucky A, Olopade OI, Cummings S, et al. Pretest prediction of *BRCA1* or *BRCA2* mutation by risk counselors and the computer model BRCAPRO. J Natl Cancer Inst. 2002 Jun 5;94(11):844-51.
181. Antoniou AC, Durocher F, Smith P, Simard J, Easton DF. *BRCA1* and *BRCA2* mutation predictions using the BOADICEA and BRCAPRO models and penetrance estimation in high-risk French-Canadian families. Breast Cancer Res. 2006;8(1):R3.
182. Katki HA. Incorporating medical interventions into carrier probability estimation for genetic counseling. BMC Med Genet. 2007;8:13.
183. Mazzola E, Blackford A, Parmigiani G, Biswas S. Recent Enhancements to the Genetic Risk Prediction Model BRCAPRO. Cancer Inform. 2015;14(Suppl 2):147-57.
184. Eisen A, Lubinski J, Klijn J, Moller P, Lynch HT, Offit K, et al. Breast cancer risk following bilateral oophorectomy in *BRCA1* and *BRCA2* mutation carriers: an international case-control study. J Clin Oncol. 2005 Oct 20;23(30):7491-6.
185. Rebbeck TR, Lynch HT, Neuhausen SL, Narod SA, Van't Veer L, Garber JE, et al. Prophylactic oophorectomy in carriers of *BRCA1* or *BRCA2* mutations. N Engl J Med. 2002 May 23;346(21):1616-22.
186. Atchley DP, Albarracin CT, Lopez A, Valero V, Amos CI, Gonzalez-Angulo AM, et al. Clinical and pathologic characteristics of patients with *BRCA*-positive and *BRCA*-negative breast cancer. J Clin Oncol. 2008 Sep 10;26(26):4282-8.
187. Biswas S, Tankhiwale N, Blackford A, Barrera AM, Ready K, Lu K, et al. Assessing the added value of breast tumor markers in genetic risk prediction model BRCAPRO. Breast Cancer Res Treat. 2012 May;133(1):347-55.

188. Mazzola E, Chipman J, Cheng SC, Parmigiani G. Recent BRCAPRO upgrades significantly improve calibration. Cancer Epidemiol Biomarkers Prev. 2014 Aug;23(8):1689-95.
189. Chen S, Wang W, Broman KW, Katki HA, Parmigiani G. BayesMendel: an R environment for Mendelian risk prediction. Stat Appl Genet Mol Biol. 2004;3:Article21.
190. Antoniou AC, Pharoah PD, McMullan G, Day NE, Stratton MR, Peto J, et al. A comprehensive model for familial breast cancer incorporating *BRCA1, BRCA2* and other genes. Br J Cancer. 2002 Jan 7;86(1):76-83.
191. Antoniou AC, Pharoah PP, Smith P, Easton DF. The BOADICEA model of genetic susceptibility to breast and ovarian cancer. Br J Cancer. 2004 Oct 18;91(8):1580-90.
192. Peto J, Mack TM. High constant incidence in twins and other relatives of women with breast cancer. Nat Genet. 2000 Dec;26(4):411-4.
193. Barcenas CH, Hosain GM, Arun B, Zong J, Zhou X, Chen J, et al. Assessing *BRCA* carrier probabilities in extended families. J Clin Oncol. 2006 Jan 20;24(3):354-60.
194. Stahlbom AK, Johansson H, Liljegren A, von Wachenfeldt A, Arver B. Evaluation of the BOADICEA risk assessment model in women with a family history of breast cancer. Fam Cancer. 2012 Mar;11(1):33-40.
195. Antoniou AC, Cunningham AP, Peto J, Evans DG, Lalloo F, Narod SA, et al. The BOADICEA model of genetic susceptibility to breast and ovarian cancers: updates and extensions. Br J Cancer. 2008 Apr 22;98(8):1457-66.
196. Lee AJ, Cunningham AP, Kuchenbaecker KB, Mavaddat N, Easton DF, Antoniou AC. BOADICEA breast cancer risk prediction model: updates to cancer incidences, tumour pathology and web interface. Br J Cancer. 2014 Jan 21;110(2):535-45.
197. MacInnis RJ, Bickerstaffe A, Apicella C, Dite GS, Dowty JG, Aujard K, et al. Prospective validation of the breast cancer risk prediction model BOADICEA and a batch-mode version BOADICEACentre. Br J Cancer. 2013 Sep 3;109(5):1296-301.
198. van Asperen CJ, Jacobi CE, de Bock GH. Remarks on the BOADICEA model of genetic susceptibility to breast and ovarian cancer. Br J Cancer. 2005 Apr 11;92(7):1336-7; author reply 7-8.
199. 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.
200. Ford D, Easton DF, Stratton M, Narod S, Goldgar D, Devilee P, et al. Genetic heterogeneity and penetrance analysis of the *BRCA1* and *BRCA2* genes in breast cancer families. The Breast Cancer Linkage Consortium. Am J Hum Genet. 1998 Mar;62(3):676-89.
201. Anderson H, Bladstrom A, Olsson H, Moller TR. Familial breast and ovarian cancer: a Swedish population-based register study. Am J Epidemiol. 2000 Dec 15;152(12):1154-63.
202. Fischer C, Kuchenbacker K, Engel C, Zachariae S, Rhiem K, Meindl A, et al. Evaluating the performance of the breast cancer genetic risk models BOADICEA, IBIS, BRCAPRO and Claus for predicting *BRCA1/2* mutation carrier probabilities: a study based on 7352 families from the German Hereditary Breast and Ovarian Cancer Consortium. J Med Genet. 2013 Jun;50(6):360-7.
203. Panchal SM, Ennis M, Canon S, Bordeleau LJ. Selecting a *BRCA* risk assessment model for use in a familial cancer clinic. BMC Med Genet. 2008;9:116.
204. Laitman Y, Simeonov M, Keinan-Boker L, Liphshitz I, Friedman E. Breast cancer risk prediction accuracy in Jewish Israeli high-risk women using the BOADICEA and IBIS risk models. Genet Res (Camb). 2013 Dec;95(6):174-7.
205. Boughey JC, Hartmann LC, Anderson SS, Degnim AC, Vierkant RA, Reynolds CA, et al. Evaluation of the Tyrer-Cuzick (International Breast Cancer Intervention Study) model for breast cancer risk prediction in women with atypical hyperplasia. J Clin Oncol. 2010 Aug 1;28(22):3591-6.

206. Evans DG, Eccles DM, Rahman N, Young K, Bulman M, Amir E, et al. A new scoring system for the chances of identifying a *BRCA1/2* mutation outperforms existing models including BRCAPRO. J Med Genet. 2004 Jun;41(6):474-80.
207. Evans DG, Lalloo F, Cramer A, Jones EA, Knox F, Amir E, et al. Addition of pathology and biomarker information significantly improves the performance of the Manchester scoring system for *BRCA1* and *BRCA2* testing. J Med Genet. 2009 Dec;46(12):811-7.
208. Kast K, Schmutzler RK, Rhiem K, Kiechle M, Fischer C, Niederacher D, et al. Validation of the Manchester scoring system for predicting *BRCA1/2* mutations in 9,390 families suspected of having hereditary breast and ovarian cancer. Int J Cancer. 2014 Nov 15;135(10):2352-61.
209. Kang HH, Williams R, Leary J, Ringland C, Kirk J, Ward R. Evaluation of models to predict *BRCA* germline mutations. Br J Cancer. 2006 Oct 9;95(7):914-20.
210. The Institute of Cancer Research. Protocol 2 (10% threshold) *BRCA1* and *BRCA2* mutation testing guidelines. [cited 2015 August]; Available from: http://d1ijoxngr27nfi.cloudfront.net/research-divisions/protocol-2-(10-threshold)-brca-mutation-testing_20150209-v1.pdf?sfvrsn=2
211. National Institute for Health and Care Excellence. Clinical evidence review. Familial breast cancer: Classification and care of women at risk of familial breast cancer and management of breast cancer and related risks in people with a family history of breast cancer. . [Online].NICE [cited 2015 August]; Available from: https://www.nice.org.uk/guidance/cg164/evidence/clinical-evidence-review-190130943
212. National Institute for Health and Care Excellence. Familial Breast Cancer: The Classification and Care of Women at Risk of Familial Breast Cancer in Primary, Secondary and Tertiary Care: Update. 2006 Jul.
213. Evans DG, Brentnall AR, Harvie M, Dawe S, Sergeant JC, Stavrinos P, et al. Breast cancer risk in young women in the national breast screening programme: implications for applying NICE guidelines for additional screening and chemoprevention. Cancer Prev Res (Phila). 2014 Oct;7(10):993-1001.
214. Maurice A, Evans DG, Shenton A, Ashcroft L, Baildam A, Barr L, et al. Screening younger women with a family history of breast cancer--does early detection improve outcome? Eur J Cancer. 2006 Jul;42(10):1385-90.
215. FH01 Collaborative Teams. Mammographic surveillance in women younger than 50 years who have a family history of breast cancer: tumour characteristics and projected effect on mortality in the prospective, single-arm, FH01 study. Lancet Oncol. 2010 Dec;11(12):1127-34.
216. The Institute of Cancer Research. Protocol 2 (10% threshold) *BRCA1* and *BRCA2* mutation testing guidelines, Frequently asked questions. [cited 2015 October]; Available from: http://d1ijoxngr27nfi.cloudfront.net/research-divisions/protocol-2-(10-threshold)-faq_20150209.pdf?sfvrsn=4
217. Groen EJ, Elshof LE, Visser LL, Rutgers EJ, Winter-Warnars HA, Lips EH, et al. Finding the balance between over- and under-treatment of ductal carcinoma in situ (DCIS). Breast. 2017 Feb;31:274-83.
218. Narod SA, Iqbal J, Giannakeas V, Sopik V, Sun P. Breast Cancer Mortality After a Diagnosis of Ductal Carcinoma In Situ. JAMA Oncol. 2015 Oct;1(7):888-96.
219. Cowell CF, Weigelt B, Sakr RA, Ng CK, Hicks J, King TA, et al. Progression from ductal carcinoma in situ to invasive breast cancer: revisited. Mol Oncol. 2013 Oct;7(5):859-69.
220. Durham TA, Koch GG, LaVange LM. Introductory Statistics in Medical Research. In: Supino PG BJ, editor. Principles of Research Methodology. New York: Springer; 2012. p. 207-31.
221. IBM Corporation. IBM SPSS Statistics for Windows. 22 ed. Armonk, New York: IBM Corporation; 2013.

222. Jones SR, Carley S, Harrison M. An introduction to power and sample size estimation. Emerg Med J. 2003 Sep;20(5):453-8.
223. Scotland ISD. Cancer Statistics, Breast Cancer. [cited 2017 January]; Available from: http://www.isdscotland.org/Health-Topics/Cancer/Cancer-Statistics/Breast
224. Gore SM, Altman DG. How large a sample. Statistics in practice. London: BMJ Publishing; 2001. p. 6-8.
225. Ziogas A, Anton-Culver H. Validation of family history data in cancer family registries. Am J Prev Med. 2003 Feb;24(2):190-8.
226. Parent ME, Ghadirian P, Lacroix A, Perret C. The reliability of recollections of family history: implications for the medical provider. J Cancer Educ. 1997 Summer;12(2):114-20.
227. Gargiulo P, Pensabene M, Milano M, Arpino G, Giuliano M, Forestieri V, et al. Long-term survival and *BRCA* status in male breast cancer: a retrospective single-center analysis. BMC Cancer. 2016 Jul 04;16:375.
228. National Institute for Health and Care Excellence. The guidelines manual PMG6. [cited February 2017, ; Available from: https://www.nice.org.uk/process/pmg6/chapter/assessing-cost-effectiveness
229. Public Health England. Critera for appraising the viability, effectiveness and appropriateness of a screening programme. . [cited 2017 January]; Available from: https://www.gov.uk/government/publications/evidence-review-criteria-national-screening-programmes

7.0. Appendices

7.1. Appendix 1: Caldicott approval Letter for study

Information Governance
Maryfield House South Mains Loan
Dundee
DD4 7BT
T. 01382 740074 Ext. 70249
www.nhstayside.scot.nhs.uk

Ms Lucy Littlejohn Medical Student
25 Thomson Street
Dundee
DD1 4LE

Date	09 January 2017
Your Ref	
Our Ref	Caldicott/CSAppLL2349++++
Enquiries	To Sender
Extension	70249
Direct	
Email	joseph.donnelly@nhs.net

Dear Ms Littlejohn

Caldicott Approval – Assessing current models of risk stratification for familial breast cancer

Proposal Sponsor:	Dr Jonathan Berg, Senior Lecturer and Honorary Consultant in Clinical Genetics, NHS Tayside
Data Users:	Dr Jonathan Berg, Senior Lecturer and Honorary Consultant in Clinical Genetics, NHS Tayside Ms Lucy Littlejohn, Medical Student, University of Dundee Subbra Palaniappan, Dr Lee Jordan, Consultant Pathologist/Honorary Senior Clinical Lecturer, Lead Clinician for the Scottish Pathology Network (SPAN), CMO's Specialty Adviser for Histopathology, NHS Tayside Dr Sarah Vinnecombe,

Original Approval 25 September 2015

Caldicott approval is given for you to extract patient data pertaining to up to 1500 women who attended clinical genetics regarding familial breast cancer risk between 2000 and 2010 and to assess, according to NICE guidelines, their risk of breast cancer at presentation. The TRACES database will then be used to identify women within the cohort who subsequently developed breast cancer in order to assess the accuracy of NICE risk prediction. The ICE system and HAMAM database will also be accessed in order to assess risk prediction models other than NICE guidelines, as described in your application and supporting information.

It is noted that all data will be pseudonymised following initial patient matching.

Extended Approval 10 February 2016

Lucy Littlejohn, Medical Student, University of Dundee included as additional data user.

Extended Approval 16 February 2016

Lyndsay McGregor, Medical Student, University of Dundee included as additional data user.

Extended Approval 06 April 2016

To allow access to Labcentre system data under the supervision of Dr Jim Gibbs, Pathology Data IT Manager, to extract breast cancer diagnosis where present and date of diagnosis.

Extended Approval 13 January 2017

Clarification that data relating to all women in Tayside who developed breast cancer under 50 may be extracted from Labcentre (not only those who attended Clinical Genetics.

Thank you for your co-operation in providing us with the information requested by us in this process. Please contact me should any queries arise from the application of this approval.

Joe Donnelly
CHI Administrator
Information Governance

Copy to:
Dr Jonathan Berg, Senior Lecturer and Honorary Consultant in Clinical Genetics, NHS Tayside Subbra Palaniappan,
Dr Lee Jordan, Consultant Pathologist/Honorary Senior Clinical Lecturer, Lead Clinician for the Scottish Pathology Network (SPAN), CMO's Specialty Adviser for Histopathology, NHS Tayside Dr Sarah Vinnecombe
Lyndsay McGregor, Medical Student, University of Dundee

7.2. Appendix 2: Clinical variables collected for each patient

<u>Patient information</u> Name Community Health Index number Date of referral to clinical genetics Age at presentation <u>History of FDRs (FDR)</u> Number of breast cancers among FDRs Mother affected (yes/no) Age mother affected (years) Sister(s) affected (yes/no) Number of breast cancers among sisters Age sister(s) affected (years) Total number of sisters Daughter(S) affected (yes/no) Number of breast cancers among daughter(s) Age daughter(s) affected (years) Total number of daughters <u>History in SDRs (SDR)</u> Number of breast cancers among SDRs Maternal aunt(s) affected (yes/no) Number of breast cancers among maternal aunts Age maternal aunt(s) affected (years) Total number of maternal aunt Paternal aunt(s) affected (yes/no) Number of breast cancers among paternal aunts Age paternal aunt (s) affected (years) Total number of paternal aunts Maternal grandmother affected (yes/no) Age maternal grandmother affected (years) Paternal grandmother affected (yes/no) Age paternal grandmother affected (years)	Maternal half-sister(s) affected (yes/no) Number of breast cancers among maternal half-sisters Age maternal half-sister(s) affected (years) Total number of maternal half-sisters Paternal half-sister(s) affected (yes/no) Number of breast cancers among paternal half-sisters Age paternal half-sister(s) affected (years) Total number of paternal half-sisters Niece(s) affected (yes/no) Number of breast cancers among nieces Age niece(s) affected (years) Total number of nieces <u>Bilateral breast cancer</u> FDR with bilateral breast cancer (yes/no) SDR with bilateral breast cancer (yes/no) Number of relatives with bilateral breast cancer Age at first breast cancer (years) Age at contralateral breast cancer (years) <u>Other cancers</u> Incidence of ovarian cancer (yes/no) Number of relatives with ovarian cancer Age of ovarian cancer(s) (years) Incidence of male breast cancer (yes/no) Age of male breast cancer (years) Number of pancreatic cancers in family Number of prostate cancers in family

7.3. Appendix 3: List of Contributors

Jim Gibbs, Department of Pathology, Ninewells Hospital and Medical School, Dundee, Angus, United Kingdom;

Simon Ogston, Department of Population Health, Ninewells Hospital and Medical School, Dundee, Angus, United Kingdom;

Dr Lee Jordan, Department of Pathology, Ninewells Hospital and Medical School, Dundee, Angus, United Kingdom;

Christine Bell, Department of Clinical Genetics, Aberdeen Royal Infirmary, Aberdeen, United Kingdom;

Dr David Goudie, Department of Clinical Genetics, Ninewells Hospital and Medical School, Dundee, Angus, United Kingdom;

Jacqueline Dunlop, Department of Clinical Genetics, Ninewells Hospital and Medical School, Dundee, Angus, United Kingdom;

Dr Jonathan Berg, Clinical Genetics, University of Dundee, Dundee, Angus, United Kingdom; Department of Clinical Genetics, Ninewells Hospital and Medical School, Dundee, Angus, United Kingdom;