DOI:
10.1158/1055-9965.EPI-18-1102

Document Version
Peer reviewed version

Frequency of pathogenic germline variants in *CDH1*, *BRCA2*, *CHEK2*, *PALB2, BRCA1* and *TP53* in sporadic lobular breast cancer

Christos Petridis[1,2], Iteeka Arora[1], Vandna Shah[1], Charlotte L. Moss[1], Anca Mera[1], Angela Clifford[1], Cheryl Gillett[1], Sarah E Pinder[1], Ian Tomlinson[3], Rebecca Roylance[4], Michael A Simpson[2], Elinor J Sawyer[1]

1. School of Cancer and Pharmaceutical Sciences, Guy's Hospital, King's College London, London SE1 9RT, UK.
2. Medical and Molecular Genetics, Guy's Hospital, King's College London, London, SE1 9RT, UK.
3. Institute of Cancer and Genomic Sciences, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK
4. Department of Oncology, UCLH Foundation Trust, London NW1 2PG, UK

(Running Title: Pathogenic germline variants in lobular breast cancer)

Corresponding author:

Elinor Sawyer, Innovation Hub, Guy's Cancer Centre, Guy's Hospital, London SE1 9RT, UK

Telephone: +44 (0) 20 7188 7843

FAX: +44 (0) 20 7188 0919

E-mail: elinor.sawyer@kcl.ac.uk

Key Words

Germline variants*, BRCA2*, *CHEK2*, *PALB2, CDH1,* invasive lobular cancer, lobular carcinoma *in situ*

Conflict of interest

The authors declare that there is no conflict of interest.

Abstract

Background: Invasive lobular breast cancer (ILC) accounts for ~15% of invasive breast carcinomas and is commonly associated with lobular carcinoma *in situ* (LCIS). Both have been shown to have higher familial risks than the more common ductal cancers. However there is little data on the prevalence of the known high and moderate penetrance breast cancer predisposition genes in ILC. The aim of this study was to assess the frequency of germline variants in *CDH1*, *BRCA2*, *BRCA1*, *CHEK2*, *PALB2* and *TP53* in sporadic ILC and LCIS diagnosed in women aged 60 years or less.

Methods: Access Array technology (Fluidigm) was used to amplify all exons of *CDH1*, *BRCA2*, *BRCA1*, *TP53*, *CHEK2* and *PALB2* using a custom made targeted sequencing panel in 1,434 cases of ILC and 368 cases of pure LCIS together with 1,611 controls.

Results: Case-control analysis revealed an excess of pathogenic variants in *BRCA2*, *CHEK2*, *PALB2* and *CDH1* in women with ILC. *CHEK2* was the only gene that showed an association with pure LCIS (OR = 9.90, 95% CI 3.42-28.66, P = 1.4 $x10^{-5}$) with a larger effect size seen in LCIS compared to ILC (OR = 4.31, 95% CI 1.61-11.58, P = 1.7 $x10^{-3}$).

Conclusions: 11% of patients with ILC aged </= 40 years carried germline variants in known breast cancer susceptibility genes.

Impact: Women with ILC aged of 40 years or less should be offered genetic screening using a panel of genes that includes *BRCA2*, *CHEK2*, *PALB2* and *CDH1*.

Introduction

Invasive lobular breast cancer (ILC) accounts for ~15% of invasive breast cancer. Its incidence has increased, in line with the greater use of combined hormone replacement therapy until the late 1990s (1,2). ILC is commonly associated with lobular carcinoma *in situ* (LCIS), which is considered both a precursor lesion and a risk factor for invasive breast cancer. The risk of invasive cancer after LCIS is 2-11 times greater than the risk in general population resulting in a cumulative long-term rate of invasive cancer of 11-26% at 15 years, in contrast to ductal carcinoma *in situ* (DCIS) which is 20 times greater (3,4). Not all invasive disease post LCIS presents as ILC, although there is an excess of ILC. Molecular studies of co-existing invasive ductal carcinoma (IDC)/ DCIS and LCIS identified similar genomic alterations, suggesting that in some cases they may have a common clonal origin (5,6). Unlike DCIS, LCIS is also a risk factor for developing invasive cancer in the contralateral breast (7). The underlying biological cause for this is not clear.

Both ILC and LCIS have been shown to have higher familial risks than the more common ductal/no special type cancers and are more likely to be bilateral (8-10). We have previously shown that many of the low risk breast cancer predisposition loci also predispose to ILC and LCIS, with some differential effects between ILC and IDC (11). We also have identified a novel lobular-specific predisposition SNP at 7q34. However, there is little data on the prevalence of the known high and moderate penetrance breast cancer predisposition genes in lobular breast cancer, with the exception of *CDH1* as lobular carcinomas are characterised by loss of E-cadherin expression (the protein encoded by *CDH1)* through somatic alterations. Germline *CDH1* variants were initially reported in hereditary diffuse gastric cancer (HDGC) (12) and about 30% of HDGC families with a *CDH1* germline variant typically have at least one individual with ILC (13-16). However germline *CDH1* variants in women with ILC and no family history of HDGC are not common (17-20), although there is some evidence that bilateral cases of LCIS + / - ILC have a higher incidence of germline *CDH1* pathogenic variants (21).

Lobular cancers have been shown to be more frequent among *BRCA2* carriers (8.4%) than *BRCA1* carriers (2.2%) (22) and there is scanty evidence that *CHEK2* and *PALB2* variants may be associated with ILC (23,24). *BRCA1* and *TP53* are not well described in ILC.

In this study we report the frequency of rare variants in six known breast cancer predisposition genes (*CDH1*, *BRCA2*, *BRCA1*, *TP53*, *CHEK2*, and *PALB2)* in 1802 sporadic UK lobular cancers diagnosed in the UK in women aged 60 years or less.

Materials and Methods

Samples

All patients and controls gave written informed consent and the studies were conducted in accordance with the Declaration of Helsinki and were approved by the following institutional review boards: GLACIER study, MREC 06/Q1702/64; King's Health Partners breast tissue bank, NHS REC ref. 12-EE-0493.

1,434 cases of ILC (with or without synchronous LCIS) and 368 cases of pure LCIS with no invasive disease diagnosed in women aged 60 or under were included in this study, together with 1,611 controls, between 6/6/2007 – 28/08/2012 (prevalent cases). Bilateral cases were eligible including those with ILC in one breast and IDC in the contralateral breast. The majority of cases (1380 ILC, 364 LCIS) were recruited through the GLACIER study from 95 hospitals throughout the UK. This study was set up with the specific aim of investigating genetic predisposition to lobular cancer in the UK. These samples were also part of our study of low risk breast cancer loci in lobular breast cancer (11). A further 58 cases, including four pure LCIS were recruited through the King's Health Partners (KHP) breast tissue bank. Cases aged 60 or under were collected in order to enrich for cases likely to have a genetic component to their disease.

Controls were identified by asking cases at the time of recruitment to identify female non-blood relatives or friends who were willing to donate a blood sample. These healthy volunteers were only eligible if they had no personal or family (up to 2nd degree) history of breast cancer, LCIS, DCIS or benign breast disease. Controls could be of any age, although older individuals were preferred, as they had lived through many of their at-risk years.

Cases and controls donated a blood sample and were asked to complete a self-administered paper-based questionnaire on their family history at the time of recruitment.

Next-Generation Sequencing

After DNA extraction from peripheral blood, Access Array technology (Fluidigm) was used to amplify all exons of *CDH1*, *BRCA2*, *BRCA1*, *TP53*, *CHEK2* and *PALB2* using a custom made targeted sequencing panel consisting of 321 amplicons (**Supplementary Table 1**). Purified libraries were quantified using Qubit High Sensitivity Assay Kit and sequenced on a HiSeq2500 (Illumina).

Bioinformatics Analysis

Primer sites from the amplicons were trimmed using Btrim and then sequences were aligned to the reference genome (http://www.novocraft.com, GRCh37 version) using Novoalign (Gap opening penalty = 65 and gap extension penalty = 7 thresholds were applied). Picard tools (v1.74 https://github.com/broadinstitute/picard) and Bedtools (v2.17.0) were used to assess coverage. Variant calling was performed using Samtools and annotated using the Annovar tool (25). We optimised the calling based on a set of variants that were positive controls and samtools was the optimal caller compared to HaplotypeCaller from GATK. The transcript that was used for each gene is reported in **Supplementary Table 1**. The frequency of

variant alleles from European reference populations was obtained from three sources (1000 genomes, ESP, ExAC).

Variants were further filtered based on read depth, quality score, and genotypic quality. All variants with a read depth < 10, quality score < 20, or genotypic score < 20 were excluded from the analysis. In addition, variants with an allelic ratio < 0.2 were excluded irrespective of read depth and variants with an allelic ratio < 0.3 and read depth < 50 were also removed.

Variants that had been previously clinically evaluated and reposited in the ClinVar database (https://www.ncbi.nlm.nih.gov/clinvar/) , were assigned labels of benign, variants of unknown significance (VUS), conflicting or pathogenic as per ClinVar.

Variants not present in the ClinVar database were considered pathogenic if they were predicted to lead to protein truncation (frameshift indels, stop-gain, stop-loss and splicing variants within 2 base pairs of the splicing junction, **Supplementary Table 2**) and VUS if they were nonsynonymous substitutions or in-frame indels.

Statistical analysis

Fisher's exact test was used for gene based rare variant analysis for both case control and case only analyses. One sided test was selected since the expectation was enrichment rather than deficit of variants in cases over controls. No adjustments have been made to account for multiple testing. With the current sample size of ILC and controls we have more than 80% power (alpha =0.05) to detect variants of combined allele frequency =0.001 and an effect size of OR = 5.

Validation

All putative pathogenic alleles and *CHEK2* 1100delC variants, identified by the above methods were confirmed by Sanger sequencing, **Supplementary Figure1**.

Results

1,434 cases of ILC (with or without synchronous LCIS), 368 cases of pure LCIS (with no invasive disease) together with 1,611 controls were analysed. The median age of cases was 52 years (interquartile range 9) and of controls was 52 (interquartile range 12), **Supplementary Table 3**. The mean coverage of our target region was 800 reads across all samples, with an average of at least 40 reads for more than 98% of the target region per sample. Of the 321 amplicons analysed seven failed to amplify consistently across the majority of the samples, however even for these seven the majority of samples had at least 10 reads for 90% of the amplicon, **Supplementary Table 4**. There was one novel pathogenic variant (PALB2:exon5:c.2487delG) that did not validate by Sanger Sequencing as there was no further DNA available.

Case-control analysis of the 1,434 cases of ILC and 1,611 controls revealed an association with putative pathogenic variants in *BRCA2*, *CDH1*, *CHEK2* and *PALB2.* No association of putative pathogenic variants was observed in *BRCA1* or *TP53*, **Table 1, Supplementary Table 5**.

BRCA2

The strongest evidence of association was with *BRCA2,* which contained 27 putative pathogenic variants in 1,434 cases and two in 1,611 controls (OR = 15.44, 95% CI 3.66-65.04, $P = 1.7.0 \times 10^{-7}$). Variants were spread throughout the gene, **Figure 1a**. Five were novel (**Supplementary Table 6**) two located in exon 11, one in exon 22 and two in exon 27. The novel frameshifts in exon 27 although unlikely to result in loss of function through nonsense mediated decay may have substantial impact on the protein product and were included as putative pathogenic variants. The two controls with pathogenic *BRCA2* variants were recruited at the age of 41 and 43 years; both were in exon 11 and have been described previously: rs80359550: c.5946delT: p.S1982fs and rs397507634: c.C2612A: p.S871X.

Case only analysis showed that *BRCA2* variants were more common in younger patients; 7% of women with ILC =/< 40 years of age were carriers and 3.4% of women =/< 50 years (case only analysis: =/< 50 years vs > 50 years: OR = 4.83, 95% CI 1.95-12.01, *P* = 0.0003, **Table 2a**). The majority (16/27; 59%) of carriers presented with ILC between 40-50 years of age. There was a borderline association with 1st degree family history of breast cancer (case only analysis OR = 2.36, 95% CI 1.05-5.31, *P* = 0.043) and any family history of breast cancer (OR = 2.25, 95% CI 1.02-4.98, *P* = 0.046), **Table 2b**. One case of pure LCIS had a *BRCA2* stop-gain variant in exon 25, c.C9294G:p.Y3098X. Case only analysis showed a significant association of *BRCA2* pathogenic variants with ILC compared to LCIS (P=0.03, **Table 1**)

CHEK2

Nineteen cases of ILC and five controls, OR = 4.31, 95% CI 1.61-11.58, *P* = 0.0017, **Supplementary Table 7**, had pathogenic variants in *CHEK2*. Seventeen cases had known pathogenic variants, two with the recently described Norwegian variant: c.C283T:p.R95X, (both bilateral); two at c.349A>G (p.Arg117Gly) and the majority (N = 13) being the well described 1100delC variant. A novel frameshift variant in exon 12 (c.1262delT:p.L421fs) was found in two cases. Four controls also had the 1100delC variant and one a splicing variant in exon15:(c.1462-2A>G). No pathogenic variants were found in women under the age of 40 and there was no association with age (case only analysis: =/< 50 years vs > 50 years: OR = 0.78, 95% CI 0.31-1.99, *P* = 0.65) or 1st degree family history of breast cancer (case only analysis OR = 2.59, 95% CI 1.02-6.60, *P* = 0.067), **Table 2b**. However, a significant excess of variants was observed in individuals with any family history of breast cancer (OR = 3.95, 95% CI 1.42-11.01, *P* = 0.008).

There was also an association with pure LCIS (OR = 9.90, 95% CI 3.42-28.66, P = 1.4 $x10^{-5}$), **Table 1**. Eleven cases had *CHEK2* variants of which eight were the 1100delC variant, one was the c.C283T:p.R95X variant and there were two novel variants: a frame shift in

exon 2 (c.188_189insC:p.L63fs) and a stop gain in exon 6 (c.G697T:p.E233X). Case only analysis revealed that *CHEK2* variants were more strongly associated with LCIS then ILC (ILC vs LCIS, P = 0.037, **Table 1**). Assessment of the pathology reports revealed that 17/19 ILC cases with pathogenic *CHEK2* variants had associated LCIS identified by the pathologists. Combining the ILC and LCIS cases gives an overall OR of 5.42 (95% CI 2.10-13.99, $P = 6x10^{-5}$) for lobular breast cancer.

PALB2

Eleven cases of ILC had pathogenic *PALB2* variants, of which seven were novel (OR = 12.45, 95% CI1.60-95.52, *P* = 0.002), **Supplementary Table 8**. One control also carried a novel frameshift variant (exon5:c.2050delC:p.P684fs). Like *CHEK2* there was no evidence of association with age (all cases were over 40 years old at diagnosis and six were over 50 (case only analysis: =/< 50 years vs > 50 years OR = 1.12 95% CI 0.34-3.68, *P* = 0.99), or with family history of breast cancer, **Table 2**. There was one case of pure LCIS with a germline *PALB2* variant in exon 10, c.G3113A:p.W1038X.

CDH1

Five cases of ILC had pathogenic *CDH1* variants, of which one was novel, and none were found in the controls (*P* = 0.02), **Supplementary Table 9**. There was an association with age (case only analysis: =/< 40 years vs >40 years, OR = 13.14, 95% CI 2.19-78.75, *P* = 0.02) but none with family history (case only with a first degree relative with breast cancer, OR = 2.95, 95% CI 0.49 -17.73, *P* = 0.24 or with any family history of breast cancer OR = 2.11, 95% CI 0.35-12.65, *P* = 0.4) **Table 2b**. There was one variant found in a case of pure LCIS, and analysis of the pathology reports showed that all the cases except one had evidence of synchronous LCIS with the ILC and in four that the LCIS was bilateral, as described previously (21), **Supplementary Table 9**.

BRCA1 and TP53

There was no evidence of an association with *BRCA1*, with just one frameshift deletion detected in exon 10 (c.2680_2681del:p.K894fs) in a patient aged 38 with bilateral LCIS, and ILC one breast and IDC and DCIS in the other, all ER positive.

There were no *TP53* variants detected in cases or controls.

Variants of unknown significance

There was no evidence of an excess of variants of unknown significance (VUS) in any of the genes including rs35187787, a rare polymorphism in *CDH1* previously associated with breast cancer (OR = 1.23, 95% CI 0.67-2.27, *P* = 0.54 for ILC), **Supplementary Table 10**. One variant in *CHEK2* classified as having conflicting interpretations of pathogenicity, rs77130927 (exon4:c.C538T:p.R180C), was found in 7 ILC cases and 1 control and showed a borderline association with ILC, (OR = 7.89, 95% CI 0.97-64.27,P = 0.05).

Bilateral Disease

The cohort included 61 cases with pure bilateral lobular cancer of which 11.5% had pathogenic variants, four in *CDH1*, two in *CHEK2* and one in *BRCA2*. There were 47 cases with ILC/LCIS in one breast and IDC/mixed invasive /DCIS in the contralateral breast, of which 8.5% had a pathogenic variant, two in *BRCA2*, one in *CHEK2* and one in *BRCA1*.

Discussion

This is the first analysis of these six known breast cancer genes in an unselected population of ILC and LCIS and unlike other studies we have evaluated the entire coding region of each gene rather than focussing on a selected number of population specific variants.

Although germline *CDH1* variants have been extensively investigated in ILC, our study shows that the majority of rare germline variants found in women with ILC occur in *BRCA2*, *PALB2*, and *CHEK2*.

A previous study has shown that ILC does occur in 8.4% of *BRCA2* variant carriers (22) none have assessed the frequency of *BRCA2* variants in a large cohort of cases purely of this special type. Our study shows that *BRCA2* variants are the most common pathogenic variants found in ILC, with 7% of ILC cases under the age of 40 carrying a *BRCA2* variant and 3% under the age of 50. The variants were distributed throughout the gene with no evidence of a lobular predisposition locus within the gene. Of the previously described variants only one was a known founder variant, 5946delT (found in two ILC cases), which has previously been associated with pancreatic and ovarian cancer (26). The known pathogenic missense variant D2723H was found in two women with ILC and has been shown to decrease *RAD51* nuclear retention even when wild-type *BRCA2* is present (27).

The most frequent *CHEK2* variant in European populations is the truncating variant c.1100delC and this has been shown to confer a two fold increase in the risk of breast cancer (28,29). There is also a rare missense variant p.I157T (c.470 T > C, rs17879961) found in 0.005% of Non-Finnish European populations which is associated with a 1.4-fold elevation in breast cancer risk (30); this is classified as a risk allele rather than a pathogenic variant and thus is not included in the list of detected pathogenic *CHEK2* variants in **Supplementary Table 7**. However this variant showed no association with ILC in our data and was detected in 3 controls and 2 cases.

Two other truncating founder variants have been described in *CHEK2,* mainly in Polish populations*:* IVS2+1G>A and del5395 (29). The 5,395 bp founder deletion that removes exons 9 and 10 of the *CHEK2* gene would not be detected using our technique, however the IVS2+1G>A, variant would be detected and was not present in this dataset.

In previous publications p.I157T and c.1100delC carriers have been reported to be associated with phenotypically different types of breast cancer, with ILC being more

common in p.I157T carriers (31-33), and IDC in c.1100delC carriers (34). In our data p.I157T was not associated with lobular breast cancer and the most common variant found in *CHEK2* was the c.1100delC variant, found in 1.3% of ILC cases in this study, a rate similar to the Breast Cancer Association Consortium (BCAC) study of 1100delC carriers, 1.2% in 4,349 lobular cases (35). The association with p.I157T and lobular histology was originally identified from Polish series and subsequently in a large meta-analysis of p.I157T performed by BCAC where the majority (93%) of the cases again came from Germany/Poland/Finland /Sweden (31), suggesting this maybe a population specific finding.

Due to the small number of pure LCIS we had limited power to detect associations with LCIS compared to ILC. *CHEK2* was the only gene in our study where pathogenic variants were also associated with pure LCIS, particularly c.1100delC, with 3% of pure LCIS cases carrying a pathogenic variant. This has not been described in previous studies of the c.1100delC variant (32-34) although the 2016 BCAC study did show a greater association with *in situ* (mostly DCIS) tumors (35). The finding that *CHEK2* variants were more strongly associated with pure LCIS than ILC suggests that *CHEK2* variants maybe predisposing to the *in situ* stage of lobular cancer with not all progressing to the invasive stage. Of the 368 pure LCIS cases in this study, only 37 have follow up data (median 81 months, range 34-333) and 13 developed subsequent invasive breast cancer, 6 ipsilateral and 7 contralateral, as previously described (36). Two of the 13 pure LCIS cases that had developed subsequent invasive disease had pathogenic variants in *CHEK2* and one a VUS in *CHEK2* (rs564605612), however a larger sample size with longer follow up would be required to ascertain whether germline variants in *CHEK2* increase the risk of subsequent invasive disease after a diagnosis of pure LCIS. Germline variants in *CHEK2* could explain the finding that invasive disease following a diagnosis of pure LCIS occurs with equal frequency in either breast and that these subsequent cancers can be in the form of ILC or IDC (7).

PALB2 was initially thought to be an intermediate risk cancer predisposition gene like *CHEK2,* but more recently the risk of breast cancer has been shown to be similar to that associated with *BRCA2*; Antoniou *et al.* reported a relative risk of 9.47 (95% CI 7.16 -12.57), compared to the general population breast-cancer incidence in the United Kingdom (37), similar to that found in the present study (OR 12.37). As 70% of breast cancers associated with germline *PALB2* variants are ER positive (a similar frequency to that found in *BRCA2* variants carriers and in sporadic breast cancer) it is not surprising that we have detected *PALB2* variants in ILC. Previously, Cybulski *et al.* showed that 0.5% of ILC (7/1306) carried a *PALB2* variant, however they only screened for the two variants common in Polish populations, 172_175delTTGT and c.509_510delGA (38).

A previous study of germline *PALB2* carriers (mainly the c.3113G>A variant) with invasive disease noted a borderline association with synchronous LCIS (39). However, we identified only one individual with pure LCIS carrying a *PALB2* truncating variant, rs180177132, in our cohort of 368 LCIS cases.

Our previous study of *CDH1* in 50 bilateral LCIS/ILC showed that 8% of bilateral lobular cases had pathogenic *CDH1* variants (21). In this current study we have identified 61 cases of bilateral lobular cancer (the majority included in our previous study), and found that 11% carry pathogenic variants in *CDH1*, *CHEK2* and *BRCA2*. Although we did not detect any new pathogenic *CDH1* variants in the bilateral cases we did identify two in unilateral cases (one in ILC+LCIS and one in pure LCIS) giving a frequency of 0.1% in unilateral disease, confirming the result of other studies that germline pathogenic variants in *CDH1* do not make a large contribution to the familial risk of unilateral lobular breast cancer.

Current UK guidelines for *BRCA1* and *BRCA2* screening are based on a minimum combined probability of variant identification in 10% of cases. All women with triple negative breast cancer under 50 years of age are considered to meet this criterion however women of this

age with ER+ breast cancer are only eligible if they have a strong family history of breast cancer or are under the age of 30. Although ILC aged 40 years or less is rare (71/1434, 5% of cases in this study) we have shown that 11% of these cases carry germline variants in known breast cancer susceptibility genes, particularly *BRCA2* (7%) and *CDH1* (3%). For women 50 years of age or less there was a 6% chance of having a germline variant and for women 60 years of age or less 5%, as variants in *CHEK2* and *PALB2* were not associated with younger age. It is likely that this is an underestimate as our methods will not detect some of the large deletions that have been described in *BRCA2* (40).

The lack of an association with age for *PALB2* variants in our study is supported by the findings of Antoniou *et al.* who showed there was a constant relative risk, irrespective of age, for *PALB2* (37). However, unlike our study, Schmidt *et al* did suggest that there is a relationship with age for the *CHEK2* c.1100delC variant (35). *PALB2* also did not show an association with family history of breast cancer and *BRCA2* only a weak association unlike *CHEK2* which showed a strong association with any family history of breast cancer but not with a history of a first degree relative with breast cancer. This is similar to the reports in pancreatic cancer series where most germline variants are found in patients without a significant family history of cancer (41).

In conclusion this study has shown that *CHEK2* variants are more frequent in LCIS than ILC. Although our study is too small to yield stable estimates of associations it gives useful estimates of prevalence in the UK population, suggesting that variants in this gene predispose to the *in situ* stage of lobular breast cancer. Longer follow up would be required to ascertain whether cases of pure LCIS with pathogenic *CHEK2* variants develop invasive disease. We have also shown that variants in *BRCA2*, *PALB2*, and *CHEK2* are more common in ILC than *CDH1* variants, although *CDH1* is still the most common germline variant in bilateral lobular cancer. Finally, we have demonstrated that women aged 40 years or less with ILC have an 11% chance of having a germline pathogenic variant indicating that

they should be offered genetic screening of *BRCA2, CHEK2, PALB2, CDH1 and BRCA1* under current UK guidance. As the majority of the invasive cancers associated with pathogenic mutations in *CHEK2*, *PALB2* or *CDH1* are ER positive, an alternative to risk reducing surgery would be chemoprevention and yearly screening with MRI, particularly for carriers of pathogenic *CHEK2* variants which are associated with a lower risk of breast cancer than variants in the other genes. Our data do not support routine genetic testing for pure LCIS under 40 as the frequency of germline mutations is age independent (<60 years 3.8%, <50 years 3.3%, <40 years 4.5%), with the exception of bilateral LCIS. However, if pathogenic germline variants in *CHEK2* were found to identify a subset of LCIS more likely to develop invasive disease then genetic testing may be of value as it would identify a group of women with LCIS that would benefit most from chemoprevention.

Acknowledgements

We thank all the individuals who took part in this study and all the staff who have enabled this work to be carried out, in particular, Maria Troy, Kelly Kohut, Pat Gorman and Michele Caneppele.

This work was supported by a King's Health Partner's Research and Development Challenge Fund (R150505, E.J.Sawyer), a Breast Cancer Campaign Project Grant (2010NovPR61, E.J.Sawyer) and the National Institute for Health Research (NIHR) Biomedical Research Centre based at Guy's and St Thomas' NHS Foundation Trust and King's College London (C.Petridis & E.J.Sawyer). The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health. Sample collection was funded by a CRUK Project grant (8146, R.Roylance & E.J.Sawyer).

References:

1. Ravdin PM. Hormone replacement therapy and the increase in the incidence of invasive lobular cancer. Breast disease **2008**;30:3-8

2. Kotsopoulos J, Chen WY, Gates MA, Tworoger SS, Hankinson SE, Rosner BA. Risk factors for ductal and lobular breast cancer: results from the nurses' health study. Breast cancer research : BCR **2010**;12:R106

3. Warnberg F, Yuen J, Holmberg L. Risk of subsequent invasive breast cancer after breast carcinoma in situ. Lancet **2000**;355:724-5

4. King TA, Pilewskie M, Muhsen S, Patil S, Mautner SK, Park A*, et al.* Lobular Carcinoma in Situ: A 29-Year Longitudinal Experience Evaluating Clinicopathologic Features and Breast Cancer Risk. Journal of clinical oncology : official journal of the American Society of Clinical Oncology **2015**;33:3945-52

5. Tazaki E, Shishido-Hara Y, Mizutani N, Nomura S, Isaka H, Ito H*, et al.* Histopathologcial and clonal study of combined lobular and ductal carcinoma of the breast. Pathology international **2013**;63:297-304

6. Wagner PL, Kitabayashi N, Chen YT, Shin SJ. Clonal relationship between closely approximated low-grade ductal and lobular lesions in the breast: a molecular study of 10 cases. Am J Clin Pathol **2009**;132:871-6

7. Chuba PJ, Hamre MR, Yap J, Severson RK, Lucas D, Shamsa F*, et al.* Bilateral risk for subsequent breast cancer after lobular carcinoma-in-situ: analysis of surveillance, epidemiology, and end results data. Journal of clinical oncology : official journal of the American Society of Clinical Oncology **2005**;23:5534-41

8. Claus EB, Risch N, Thompson WD, Carter D. Relationship between Breast Histopathology and Family History of Breast-Cancer. Cancer **1993**;71:147-53

9. Cannonalbright LA, Thomas A, Goldgar DE, Gholami K, Rowe K, Jacobsen M*, et al.* Familiality of Cancer in Utah. Cancer research **1994**;54:2378-85

10. Claus EB, Risch NJ, Thompson WD. Using Age of Onset to Distinguish between Subforms of Breast-Cancer. Ann Hum Genet **1990**;54:169-77

11. Sawyer E, Roylance R, Petridis C, Brook MN, Nowinski S, Papouli E*, et al.* Genetic Predisposition to In Situ and Invasive Lobular Carcinoma of the Breast. Plos Genet **2014**;10

12. Guilford P, Hopkins J, Harraway J, McLeod M, McLeod N, Harawira P*, et al.* E-cadherin germline mutations in familial gastric cancer. Nature **1998**;392:402-5

13. Brooks-Wilson AR, Kaurah P, Suriano G, Leach S, Senz J, Grehan N*, et al.* Germline E-cadherin mutations in hereditary diffuse gastric cancer: assessment of 42 new families and review of genetic screening criteria. Journal of medical genetics **2004**;41:508-17

14. Kaurah P, MacMillan A, Boyd N, Senz J, De Luca A, Chun N*, et al.* Founder and recurrent CDH1 mutations in families with hereditary diffuse gastric cancer. JAMA : the journal of the American Medical Association **2007**;297:2360-72

15. Pharoah PD, Guilford P, Caldas C, International Gastric Cancer Linkage C. Incidence of gastric cancer and breast cancer in CDH1 (E-cadherin) mutation carriers from hereditary diffuse gastric cancer families. Gastroenterology **2001**;121:1348-53

16. Suriano G, Yew S, Ferreira P, Senz J, Kaurah P, Ford JM*, et al.* Characterization of a recurrent germ line mutation of the E-cadherin gene: implications for genetic testing and clinical management. Clinical cancer research : an official journal of the American Association for Cancer Research **2005**;11:5401-9

17. Masciari S, Larsson N, Senz J, Boyd N, Kaurah P, Kandel MJ*, et al.* Germline E-cadherin mutations in familial lobular breast cancer. Journal of medical genetics **2007**;44:726-31

18. Schrader KA, Masciari S, Boyd N, Salamanca C, Senz J, Saunders DN*, et al.* Germline mutations in CDH1 are infrequent in women with early-onset or familial lobular breast cancers. Journal of medical genetics **2011**;48:64-8

19. Xie ZM, Li LS, Laquet C, Penault-Llorca F, Uhrhammer N, Xie XM*, et al.* Germline Mutations of the E-Cadherin Gene in Families With Inherited Invasive Lobular Breast Carcinoma But No Diffuse Gastric Cancer. Cancer **2011**;117:3112-7

20. Rahman N, Stone JG, Coleman G, Gusterson B, Seal S, Marossy A*, et al.* Lobular carcinoma in situ of the breast is not caused by constitutional mutations in the E-cadherin gene. British journal of cancer **2000**;82:568-70

21. Petridis C, Shinomiya I, Kohut K, Gorman P, Caneppele M, Shah V*, et al.* Germline CDH1 mutations in bilateral lobular carcinoma in situ. British journal of cancer **2014**;110:1053-7

22. Mavaddat N, Barrowdale D, Andrulis IL, Domchek SM, Eccles D, Nevanlinna H*, et al.* Pathology of Breast and Ovarian Cancers among BRCA1 and BRCA2 Mutation Carriers: Results from the Consortium of Investigators of Modifiers of BRCA1/2 (CIMBA). Cancer Epidem Biomar **2012**;21:134-47

23. Cybulski C, Gorski B, Huzarski T, Byrski T, Gronwald J, Debniak T*, et al.* CHEK2-positive breast cancers in young Polish women. Clin Cancer Res **2006**;12:4832-5

24. Teo ZL, Park DJ, Provenzano E, Chatfield CA, Odefrey FA, Tu ND*, et al.* Prevalence of PALB2 mutations in Australasian multiple-case breast cancer families. Breast Cancer Research **2013**;15

25. Hintzsche JD, Robinson WA, Tan AC. A Survey of Computational Tools to Analyze and Interpret Whole Exome Sequencing Data. Int J Genomics **2016**;2016:7983236

26. Rebbeck TR, Mitra N, Wan F, Sinilnikova OM, Healey S, McGuffog L*, et al.* Association of type and location of BRCA1 and BRCA2 mutations with risk of breast and ovarian cancer. Jama **2015**;313:1347-61

27. Jeyasekharan AD, Liu Y, Hattori H, Pisupati V, Jonsdottir AB, Rajendra E*, et al.* A cancer-associated BRCA2 mutation reveals masked nuclear export signals controlling localization. Nature structural & molecular biology **2013**;20:1191-8

28. CHEK2*1100delC and susceptibility to breast cancer: a collaborative analysis involving 10,860 breast cancer cases and 9,065 controls from 10 studies. American journal of human genetics **2004**;74:1175-82

29. Cybulski C, Wokolorczyk D, Jakubowska A, Huzarski T, Byrski T, Gronwald J*, et al.* Risk of Breast Cancer in Women With a CHEK2 Mutation With and Without a Family History of Breast Cancer. J Clin Oncol **2011**;29:3747-52

30. Kilpivaara O, Vahteristo P, Falck J, Syrjakoski K, Eerola H, Easton D*, et al.* CHEK2 variant I157T may be associated with increased breast cancer risk. International journal of cancer **2004**;111:543-7

31. Muranen TA, Blomqvist C, Dork T, Jakubowska A, Heikkila P, Fagerholm R*, et al.* Patient survival and tumor characteristics associated with CHEK2:p.I157T - findings from the Breast Cancer Association Consortium. Breast cancer research : BCR **2016**;18:98

32. Huzarski T, Cybulski C, Domagala W, Gronwald J, Byrski T, Szwiec M*, et al.* Pathology of breast cancer in women with constitutional CHEK2 mutations. Breast cancer research and treatment **2005**;90:187-9

33. Domagala P, Wokolorczyk D, Cybulski C, Huzarski T, Lubinski J, Domagala W. Different CHEK2 germline mutations are associated with distinct immunophenotypic molecular subtypes of breast cancer. Breast cancer research and treatment **2012**;132:937-45

34. Schmidt MK, Tollenaar RA, de Kemp SR, Broeks A, Cornelisse CJ, Smit VT*, et al.* Breast cancer survival and tumor characteristics in premenopausal women carrying the CHEK2*1100delC germline mutation. Journal of clinical oncology : official journal of the American Society of Clinical Oncology **2007**;25:64-9

35. Schmidt MK, Hogervorst F, van Hien R, Cornelissen S, Broeks A, Adank MA*, et al.* Age- and Tumor Subtype-Specific Breast Cancer Risk Estimates for CHEK2*1100delC Carriers. J Clin Oncol **2016**;34:2750-60

36. Shah V, Nowinski S, Levi D, Shinomiya I, Kebaier Ep Chaabouni N, Gillett C*, et al.* PIK3CA mutations are common in lobular carcinoma in situ, but are not a biomarker of progression. Breast cancer research : BCR **2017**;19:7

37. Antoniou AC, Casadei S, Heikkinen T, Barrowdale D, Pylkas K, Roberts J*, et al.* Breast-cancer risk in families with mutations in PALB2. The New England journal of medicine **2014**;371:497-506

38. Cybulski C, Kluzniak W, Huzarski T, Wokolorczyk D, Kashyap A, Jakubowska A*, et al.* Clinical outcomes in women with breast cancer and a PALB2 mutation: a prospective cohort analysis. Lancet Oncol **2015**;16:638-44

39. Teo ZL, Provenzano E, Dite GS, Park DJ, Apicella C, Sawyer SD*, et al.* Tumour morphology predicts PALB2 germline mutation status. Brit J Cancer **2013**;109:154-63

40. McVeigh TP, Cody N, Carroll C, Duff M, Farrell M, Bradley L*, et al.* Recurrent large genomic rearrangements in BRCA1 and BRCA2 in an Irish case series. Cancer Genet **2017**;214-215:1-8

41. Shindo K, Yu J, Suenaga M, Fesharakizadeh S, Cho C, Macgregor-Das A*, et al.* Deleterious Germline Mutations in Patients With Apparently Sporadic Pancreatic Adenocarcinoma. J Clin Oncol **2017**;35:3382-90

TABLES

Table 1: Association of known pathogenic variants with ILC and LCIS in women =/< 60 years of age by gene (1,434 ILC, 368 LCIS cases, 1,611 controls).

Gene	Pathogenic Variants in ILC Cases	Pathogenic Variants in LCIS Cases	Pathogenic Variants in Controls	OR (95% CI) for ILC	P value for ILC	OR (95% CI) For LCIS	P value for LCIS	P-Value Case only (ILC vs LCIS)
BRCA2	27	1	2	15.44 (3.66-65.04)	1.7×10^{-7}	2.19 (0.20-24.24)	0.46	0.030
CHEK2	19	11	5	4.31 (1.61-11.58)	1.7×10^{-3}	9.90 (3.42-28.66)	1.4×10^{-5}	0.037
PALB2	11	1	1	12.45 (1.60-95.52)	2.2×10^{-3}	4.39 (0.27-70.30)	0.34	0.48
CDH1	5	1	0	-	0.02	-	0.18	1.00
BRCA1	1	0	0	-	0.5	-	-	-
TP53	0	0	0	-	-	-	-	-

Table 2a: Case only analysis of pathogenic variants in ILC by age.

	Carriers </=40 vs >40		Carriers </=50 vs 51-60	
Gene	**OR (95% CI)**	**P**	**OR (95% CI)**	**P**
BRCA2	4.62 (1.74-12.23)	0.009	4.83 (1.95-12.01)	0.0003
CHEK2	0.00 (0.00-0.00)	0.62	0.78 (0.31-1.99)	0.65
PALB2	0.00 (0.00-0.00)	1	1.12 (0.34-3.68)	1
CDH1	13.14 (2.19-78.75)	0.022	2.02 (0.34-12.1)	0.66
BRCA1	-	0.05	-	0.43
TP53	-	-	-	-

Table 2b: Case only analysis of pathogenic variants in ILC for family history (FH) of breast cancer (BC)

	Carriers FH of BC vs no FH in first degree relative		Carriers FH of BC vs no FH in Any relative	
Gene	**OR (95% CI)**	**P**	**OR (95% CI)**	**P**
BRCA2	2.36 (1.05-5.31)	0.043	2.25 (1.02-4.98)	0.046
CHEK2	2.59 (1.02-6.60)	0.067	3.95 (1.42-11.01)	0.008
PALB2	2.94 (0.83-10.45)	0.097	3.26 (0.84-12.65)	0.105
CDH1	2.95 (0.49-17.73)	0.238	2.11 (0.35-12.65)	0.415
BRCA1	0	1	-	0.419
TP53	-	-	-	-

FIGURE LEGEND

Figure 1: Position of pathogenic variants

Figure 1 shows the position of pathogenic variants in (a) *BRCA2*, (b) *CHEK2*, (c) *PALB2*, (d) *CDH1* (black = case, green = control)