phgfoundation
making science work for health

Whole genome sequencing for breast cancer risk testing

July 2016

Authors:

Dr Charlotte Warren-Gash, Dr Mark Kroese, Dr Hilary Burton, Professor Paul Pharoah

NB: URLs in this report were correct as of 1 April 2016

This report can be downloaded from our website:
www.phgfoundation.org

Published by PHG Foundation
2 Worts Causeway
Cambridge
CB1 8RN
UK

Tel: +44 (0) 1223 761 900

Correspondence to:

mark.kroese@phgfoundation.org

How to reference this report:

Whole genome sequencing in screening for breast cancer (2016)
ISBN 978-1-907198-21-2

Contents

Executive summary

This work investigates the impact of routine testing for inherited / heritable conditions using whole genome sequencing where these conditions are not part of the primary purpose of testing, using the example of inherited breast cancer.

Genetic testing can be used to identify variants associated with increased disease risk

Since the completion of the Human Genome Project, the cost of whole genome sequencing has dropped enough that it is increasingly accessible to physicians, researchers and the public. Using such a technology has great potential benefit, as sequencing a whole genome may generate unlooked-for results which may be of clinical significance.

Many uncertainties impact upon the utility of using whole genome sequencing

However there is a wide range of potential psychological, ethical, legal and social implications of using WGS to test for heritable variants in patients at low prior risk. Cost implications of the technology and manpower for sequencing, interpretation and pre and post-test counselling are also a concern.

A major hurdle to implementation is difficulty interpreting the clinical relevance of genomic variants detected. There are also technical concerns about the greater accuracy needed to report findings in the clinical setting compared to the research laboratory. Every result has to be judged on whether there is sufficient evidence to support pathogenicity and thereby reduce the risk of reporting false positives.

It is therefore imperative to ensure that robust processes are in place for managing and understanding these complex data and appreciating the levels of uncertainty about clinical validity and utility of testing positive for a pathogenic variant.

Further investigation into the benefits and risks of WGS based testing for inherited breast cancer risk in unselected populations needs to be undertaken

This work demonstrates the type of process that should be undertaken when considering likely test and health outcomes that may arise when using WGS in unselected populations, using the example of inherited breast cancer variants. The aim of detecting such genetic variants is to quantify the absolute risk of disease conferred by the variant for an individual. This will allow valid judgements to be made about the risks and benefits of clinical interventions for the patient and family.

The performance of WGS for detecting inherited breast cancer variants should be monitored

Accurate knowledge of the absolute cancer risk associated with a mutation in a particular population will allow gene variants conferring risk above a certain threshold to be fed back to clinicians and patients thereby enabling appropriate clinical action. To optimise clinical diagnostic value, WGS should detect clinically actionable variants of all classes with very high sensitivity and specificity.

WGS is likely to have good potential to detect the 90% of *BRCA1* and *BRCA2* mutations that are due to small sequence variants (in particular substitutions) but may perform less well for large rearrangements and duplications, although this is a rapidly developing area. This falls short of the 95% sensitivity that is aimed for with current genetic diagnostic tests.

In our model - a hypothetical population of 100,000 UK women - WGS would correctly identify 132 women out of 173 with a pathogenic mutation and 12,633 women with breast cancer (with/without mutations). Developments in sequencing technology, software and algorithms used to report variants mean that the test performance figures outlined are likely to change over time.

Conclusions

At present there are many uncertainties that will impact on the clinical utility of actively seeking secondary findings using WGS, even for relatively well-characterised genes such as *BRCA1* and *BRCA2*. The process demonstrated in this paper when applied to other gene-disease combinations is likely to highlight further gaps in knowledge and understanding around clinical validity and utility, which should be considered before offering such routine testing of genes for secondary purposes in unselected populations.

1. Background

Aim

The aim of this work is to investigate the impact of routine testing for inherited / heritable conditions using whole genome sequencing where these conditions are not part of the primary purpose of testing, using the example of inherited breast cancer.

Objectives

- To consider which inherited breast cancer gene variants it may be appropriate to test for in unselected adult populations
- To review population prevalence and gene-disease association from the literature for each of these gene variants
- To outline factors needed to establish test performance for WGS-based testing using the examples of *BRCA1* and *BRCA2*
- To discuss positive and negative impacts of such an approach for routine clinical practice

1.1 Whole genome sequencing: an introduction

Sequencing the first human genome as part of the Human Genome Project in 2003 took over a decade and cost 3 billion USD [1]. Since then, development of next generation sequencing technologies has made whole genome sequencing a clinical reality, often achievable in as little as 24 hours for the cost of a few thousand pounds [2]. Technologies such as whole genome sequencing (WGS) (and also whole exome sequencing which is not the primary focus of this report but to which many of the same issues apply) are now increasingly accessible to physicians, researchers, patients and the public [3]. While they offer new potential to diagnose rare inherited diseases, a major hurdle to implementation is difficulty interpreting the clinical relevance of genomic variants detected [4]. Although sequencing itself may be rapid, interpretation of the large volume of complex data generated – around 3-4 million genetic variants that differ from the reference human genome – is not: turnaround times for reports of whole genome sequences from clinical laboratories tend to be in the order of >90 days [5]. This means that total costs of WGS may be considerably more than advertised [6].

Nonetheless, the potential for next generation sequencing to detect people at high genetic risk of diseases such as cancer is well recognised: a number of commercial gene panels are available to test asymptomatic individuals for cancer risk genes. Although technologies such as WGS are not at present used to screen asymptomatic people at population risk for cancer susceptibility, sequencing a whole genome may generate unexpected, or 'incidental' findings of mutations in cancer susceptibility genes. Whether or under what circumstances to return such findings to patients is the subject of intense policy debate.

1.2 Return of secondary findings from WGS: the policy

American College of Medical Genetics and Genomics (ACMG)

In March 2013 the ACMG Board approved controversial recommendations that for clinical exome or genome sequencing tests ordered for any indication, a list of 57 (later revised to 56) genes and types of variants should be routinely examined (appendix 1) and results reported to the ordering clinician, without seeking patient or family preference and regardless of patient age [7]. The list includes the genes *BRCA1* and *BRCA2*, in which certain mutations are associated with hereditary breast and ovarian cancer syndrome. These recommendations resulted from a year-long consensus process and expert consultation by an ACMG Working Group on incidental findings in clinical exome and genome sequencing. It was recognised that the evidence on disease penetrance and clinical utility was lacking to fully support the recommendations and the need for an iterative process (to take new evidence into account) was highlighted.

Following their publication, major objections to these recommendations from numerous commentators [8];[9];[10];[11];[12];[13] related to:

1. Lack of respect for patient autonomy and the right 'not to know'
2. Failure to consider patient age e.g. the implications of testing children for adult onset conditions
3. Uncertainty about the balance of benefit versus harm of interventions for high risk gene carriers without a strong family history of disease
4. Issues around informed consent
5. Technical concerns about the greater accuracy needed to report findings in the clinical setting compared to the research laboratory
6. Cost implications of technology and manpower for sequencing, interpretation and pre- and post-test counselling
7. Risk of increased health inequalities

An update to these recommendations was published by the ACMG in 2014 to allow patients to have the opportunity to opt out of analysis of medically actionable genes when undergoing whole genome sequencing [14]. Other recommendations remained unchanged.

In response to a Presidential Commission for the Study of Bioethical Issues report [15], the ACMG is also adopting the term 'secondary findings' in place of 'incidental findings'. In this report, secondary findings are defined as findings that are 'actively sought by a practitioner (but are) not the primary target (of the investigation)'. In contrast, incidental findings are unsolicited findings which might be 'anticipatable' i.e. known to be associated with a test or procedure or 'unanticipatable' i.e. unable to be anticipated given the current state of scientific knowledge.

Table 1: Categories of findings of whole genome sequencing

Type of genomic finding	Definition
Pertinent (or primary)	Pertinent findings are relevant to the explanation, main diagnosis or treatment of the disease for which the patient was referred
Secondary	Secondary findings are additional looked-for findings of healthcare importance that are not pertinent to the main condition
Incidental (or unsolicited)	Incidental findings are additional findings that are not actively sought

European Society of Human Genetics

A contrasting approach to whole genome analysis is taken by the European Society of Human Genetics which recommends using a targeted testing or reporting strategy where possible to minimise the risk of generating unsolicited findings [16]. Genomic screening is not specifically advocated and must be judged necessary to solve a diagnostic problem as well as having an acceptable balance of potential benefits versus harms for the individual being tested. Deliberately seeking out gene variants unrelated to the presenting clinical problem is discouraged. A protocol to give guidance on return of incidental findings whenever techniques of whole genome analysis are used is recommended, as are developing specific guidance for informed consent and testing of minors.

100,000 Genomes Project

In England, plans to sequence 100,000 genomes from NHS patients were announced in December 2012, with the organisation Genomics England set up by the Department of Health in July 2013 to deliver this project [17]. It aimed to focus on sequencing genomes from cancer patients, from patients with rare diseases and their close relatives, and from infectious disease agents. The project's policy on secondary findings suggests that Genomics England will look for actionable genomic findings known to cause serious conditions. A list of 16 genes related to 10 conditions, in which mutations will be actively sought is given, which includes *BRCA1* and *BRCA2* genes (Appendix 2). Participants will have the opportunity to opt out of receiving any secondary findings and testing for variants associated with adult-onset conditions will not be offered to children.

The approach taken by Genomics England in the 100,000 Genomes Project has potential to set a precedent for an NHS approach to secondary and incidental genomic findings which could have far-reaching implications for patients, their families and wider society. It is therefore imperative to ensure that robust processes are in place for managing and understanding these complex data and appreciating the levels of uncertainty about clinical validity and utility of testing positive for a cancer risk gene.

1.3 Current recommendations for detecting breast cancer risk alleles

In England, the National Institute for Health and Care Excellence (NICE) recommends that decisions to refer people to a specialist clinic for testing for high risk breast cancer gene variants should be based on family history and a carrier probability calculation method such as BOADICEA or the Manchester risk score [18]. Tests should be carried out on family members affected by cancer where possible in the first instance to try to identify a mutation in the appropriate gene (such as *BRCA1*, *BRCA2* or *TP53*). A search / screen for a mutation in a gene in order to detect coding alterations should aim for as close to 100% sensitivity as possible and the whole gene(s) should be searched.

In the US, the Preventive Services Task Force recommends that women with family histories of breast, ovarian, tubal or peritoneal cancer should be screened with screening tools and those with positive screens should receive genetic counselling and, if indicated after counselling, *BRCA* testing [19]. It recommends against routine genetic counselling or *BRCA* testing in women without a family history suggestive of increased mutation risk. This is based on lack of evidence around cancer risk for mutation carriers from the general population rather than from high risk cancer families.

There has, however, been a recent call for genetic screening to be offered to all women in the United States from the age of 30, regardless of family history, because of new evidence from Ashkenazi Jewish populations in Israel that mutation carriers without a strong family history of breast cancer are still at high risk of disease [20]. It is not clear, though, that the findings are generalizable to other populations likely to have a lower prevalence of pathogenic mutations and in whom penetrance may differ compared to women in the study.

1.4 Options for breast cancer screening

Universal screening based on age

In the UK the NHS Breast Screening Programme offers 3-yearly mammographic screening to women aged 50 to 70 [21] with the aim of detecting early breast cancers and thereby reducing mortality. Eligibility for this programme is being extended to those aged 47 to 73 years.

Enhanced screening based on cancer risk

More frequent surveillance is recommended for women who have either been diagnosed with breast cancer in the past or who have higher cancer risk for other reasons e.g. because they carry a high risk allele, with the choice of imaging modality related to a woman's age [21].

Stratified screening based on age and polygenic risk profile

It has recently been proposed that, using a combination of age and polygenic risk profile based on the published risks associated with 17 breast cancer susceptibility loci, the population could be stratified into groups with different absolute risks of breast cancer [22]. Each group would start mammographic screening at a different age threshold based on when their absolute 10-year risk of breast cancer reached 2.5% (which is the current threshold for the lower age limit of 47 years in the National Breast Screening programme). This personalised screening approach has the potential to improve the efficiency of screening programmes through more appropriate targeting of resources.

1.5 Opportunistic screening using whole genome sequencing testing

Screening for existing disease to enable early treatment should be distinguished from screening for genetic variants that place an individual at higher risk of disease to allow preventive action. For the former, the National Screening Committee criteria which include considerations around the condition, the test, treatment available and the screening programme more widely, should be met before a population screening programme is implemented [23]. Additional evidence required for screening for genetic variants can be evaluated using frameworks such as ACCE [24], in which analytical validity, clinical validity, clinical utility and ethical, legal and social implications are reviewed. Important steps to such an evaluation include considering how accurately a genetic test will detect all relevant mutations (test performance), defining disease risk in those with an affected gene (gene-disease association), ascertaining population prevalence of the gene variant (as phenotypic association may depend on prior probability of disease) and weighing relative harms and benefits of management options in carriers of high risk gene variants.

2. Results

In our model - a hypothetical population of 100,000 UK women - WGS would correctly identify 132 women out of 173 with a pathogenic mutation and 12,633 women with breast cancer (with/without mutations). Developments in sequencing technology, software and algorithms used to report variants mean that the test performance figures outlined are likely to change over time.

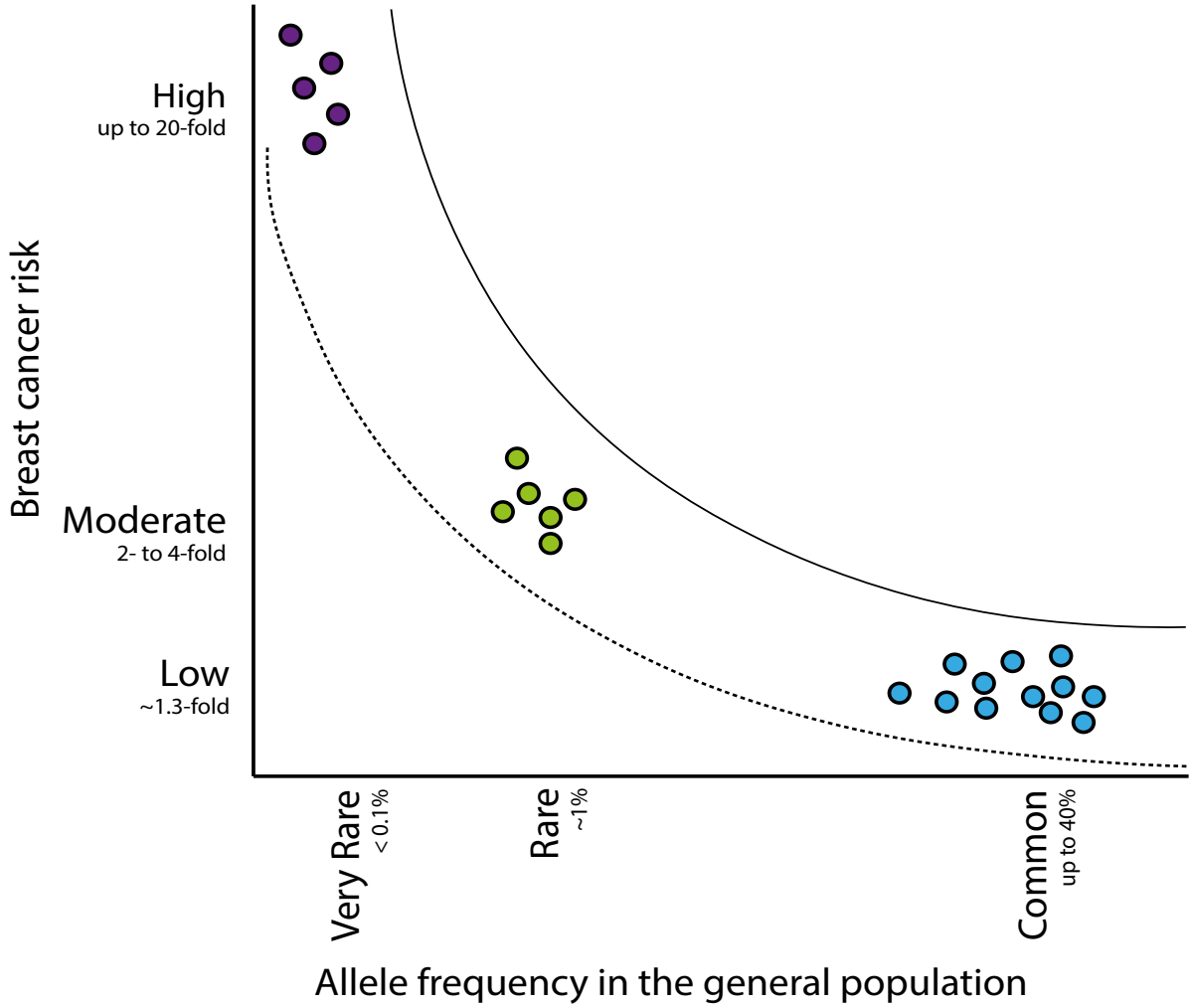
2.1 Genetic basis of inherited breast cancer

Over 100 single nucleotide polymorphisms (SNPs) associated with breast cancer risk have been identified through large genetic studies. These tend to be relatively common and each confers a relative risk of <1.5 [25]. These common variants are relatively distinct from rarer moderate risk breast cancer alleles (that confer a relative risk of two to four fold) or high risk alleles (that are associated with a >4 fold relative risk of breast cancer) such as those found in families with multiple affected members. Inherited forms of breast cancer, attributed to these moderate and high risk gene variants contribute to up to 10% of all breast cancer cases. In the UK in 2011, 49,900 women were diagnosed with breast cancer [26], so it can be estimated that just under 5,000 of these cancers were due to rare risk alleles. The incidence rate of breast cancer per year in the UK for women of all ages is 155.2/100,000 and lifetime population risk is around 12.5% [26].

The population frequency of different alleles associated with breast cancer risk versus the relative risk conferred by these alleles is shown in the figure below (Figure 1).

In a recent review of the association between genes offered in commercial panel tests and susceptibility to breast cancer, nine genes were found to have clear evidence of involvement in breast cancer susceptibility (*ATM, BRCA1, BRCA2, CHEK2, NBN, PALB2, PTEN, STK11, TP53*) [25]. Only four of these – *BRCA1*, *BRCA2*, *PALB2, TP53* – conferred a greater than four fold risk of disease such that they could be classified as 'high risk' genes. These genes are considered below.

Figure 1: Allelle frequency versus breast cancer risk. From Hollestalle *et al.* [27]



2.2 *BRCA1* and *BRCA2*

Location, function and role in cancer pathogenesis

BRCA1 and *BRCA2*, located at chromosomal loci 17q21.31 and 13q13.1 respectively, were discovered in the early 1990s and are the best known examples of breast cancer susceptibility genes [28]. They are tumour suppressor genes with an integral role in response to DNA damage and particularly double-strand break repair by homologous recombination [28]; [29]. In all cancers from individuals with a germline *BRCA1* mutation for example, the normal allele is deleted or inactivated resulting in homozygous inactivation of *BRCA1* [28]. Cells that lack a functional *BRCA1* protein accumulate chromosomal abnormalities, leading to genomic instability that underlies malignant transformation and breast cancer pathogenesis. Germline mutations in *BRCA1* and *BRCA2* account for the majority of cases of hereditary breast and ovarian cancer syndrome [30].

Over 1,600 germline pathogenic allelic variants in *BRCA1* and 1,800 in *BRCA2* have been described, with most leading to frameshifts resulting in a missing or non-functional protein [28]. The term 'protein-truncating variants' typically includes frameshift insertions or deletions (indels), single base nonsense substitutions and variants affecting splicing [25]. Approximately 88-90% of pathogenic *BRCA1* and *BRCA2* mutations are due to sequence variants including small indels and the remaining 10-12% of mutations are due to exonic or whole gene deletions or duplications [31]. This is important because the capability of a technology such as WGS to detect *BRCA* mutations will vary depending on mutation type.

The proportion of variants of uncertain significance (VUS), often missense mutations, varies depending upon the population being tested as this affects prior probability that a mutation is pathogenic. In *BRCA1* and *BRCA2* the proportion of VUS is typically around 10-15%, although both higher figures (of up to 30% in *BRCA1* and 50% in *BRCA2* [32]) have been reported as well as much lower figures – of down to 2.9% [33]. This wide spectrum reflects both the dynamic nature of the field (whereby variants are reclassified as pathogenic or benign as research generates new understanding of pathogenicity) and the substantial differences that may arise due to differing population ethnicity or family cancer history.

Population prevalence of BRCA1 and BRCA2 mutations

Table 2 shows some examples of studies investigating the population prevalence of deleterious germline variants in *BRCA1* and *BRCA2*. Prevalence varies depending on the population and study, but figures of 1 in 300 to 1 in 500 are most consistently reported for the population carrier frequency of both mutations. Founder mutations, which are observed in populations derived from a small ancestral group, have been described at much higher frequencies in diverse groups from Canada, Hungary, Iceland, Italy, Netherlands and Sweden as well as Ashkenazi Jewish populations [34]. No founder mutations have been reported in the UK.

Table 2: Prevalence of *BRCA1* and *BRCA2* mutations in various non-cancer populations

Population	Proportion of women without breast cancer who are *BRCA* mutation carrier	Allele frequency	Reference
UK			
BRCA1	1 in 833 (approx.)	0.0006	Ford 1995 [35]
Worldwide			
Same risks assumed for *BRCA1* and *BRCA2*	1 in 397 overall	0.0006 each for average risk womenα 0.0075 each for moderate risk womenβ 0.0434 each for high risk womenγ	Nelson 2005 [36]
Worldwide			
Same risks assumed for *BRCA1* and *BRCA2*	1 in 300 - 1 in 500	0.001 - 0.00167	Christinat 2013 [37]
UK			
BRCA1 *BRCA2*	1 in 833 1 in 500	0.0006 0.0010	Antoniou 2008 [38]
Ashkenazi Jews			
Any of three founder mutations in *BRCA1* or 2	1 in 40	0.0125	Christinat 2013 [37]
BRCA1 185delAG *BRCA1* 5382insC *BRCA2* 6174delT	0.8% - 1% 0.1% - 0.4% 1% - 1.5%	0.004 – 0.005 0.002 – 0.0005 0.005 – 0.0075	Fackenthal 2007 [34]
Iceland			
BRCA2 999del5	0.4%-0.6%	0.002 – 0.003	Fackenthal 2007 [34]

α Low risk assumes mutation prevalence rate of 0.12% for average risk women (no first degree relatives and no more than one second degree relative affected by breast or ovarian cancer on either side. This accounts for 92.7% of screening population)
β Medium risk assumes mutation prevalence rate of 1.5% for moderate risk women (one first degree relative or two second degree relatives affected by breast or ovarian cancer on same side. This accounts for 6.9% of screening population)
γ High risk assumes mutation prevalence rate of 8.68% for high risk women (at least two first degree relatives with breast or ovarian cancer. This accounts for 0.4% of screening population)

Penetrance of BRCA1 and BRCA2 mutations

There is no single 'penetrance' or absolute risk of breast cancer conferred by *BRCA1* and *BRCA2* mutations: absolute risk varies with age, family history and mutation type, depending on a complex interplay between the genetic variant detected and a range of environmental and other genetic modifiers. Figure 2a shows the different lifetime breast cancer risks associated with the same mutation in individuals with the same family history tested at different ages. Figure 2b shows how lifetime breast cancer risk varies in individuals with different family histories tested at the same age. The concept of intra-allelic risk heterogeneity has also been described. An example is a nonsense variant at the carboxyl terminus of *BRCA2*, K3326X, that has been associated with a relative risk of breast cancer of only 1.39 [39].

Figure 2a: Breast cancer risks associated with the same *BRCA1* mutation arising in women tested at different ages who have the same family history (based on BOADICEA risk calculations 3 and 10)

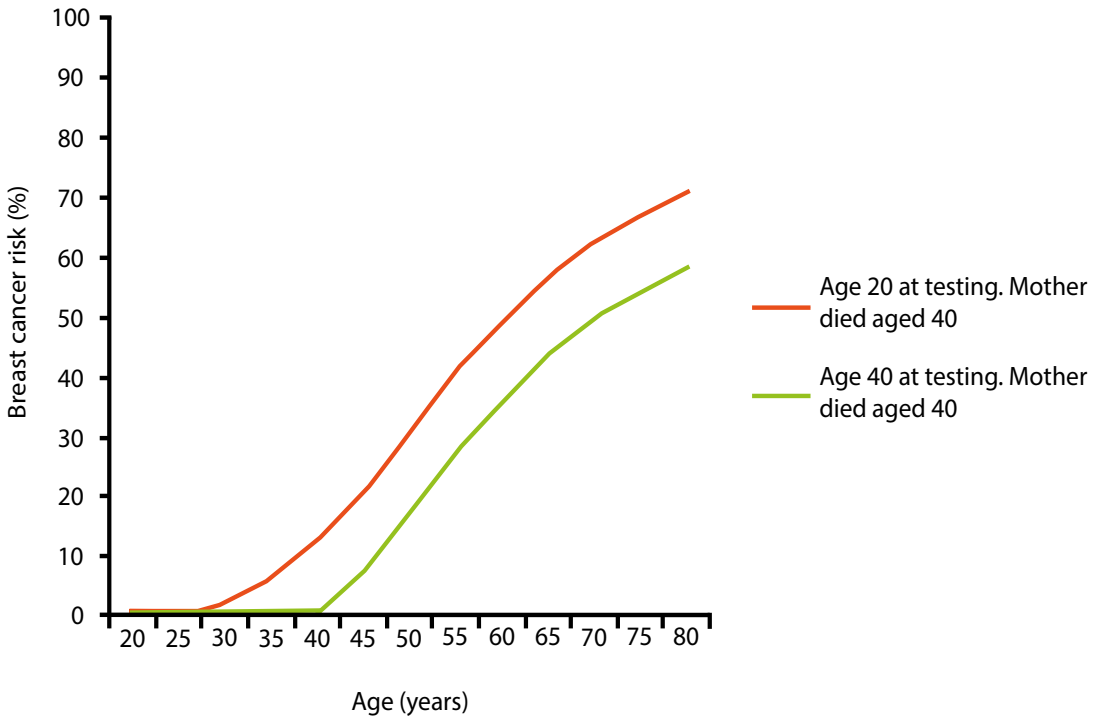
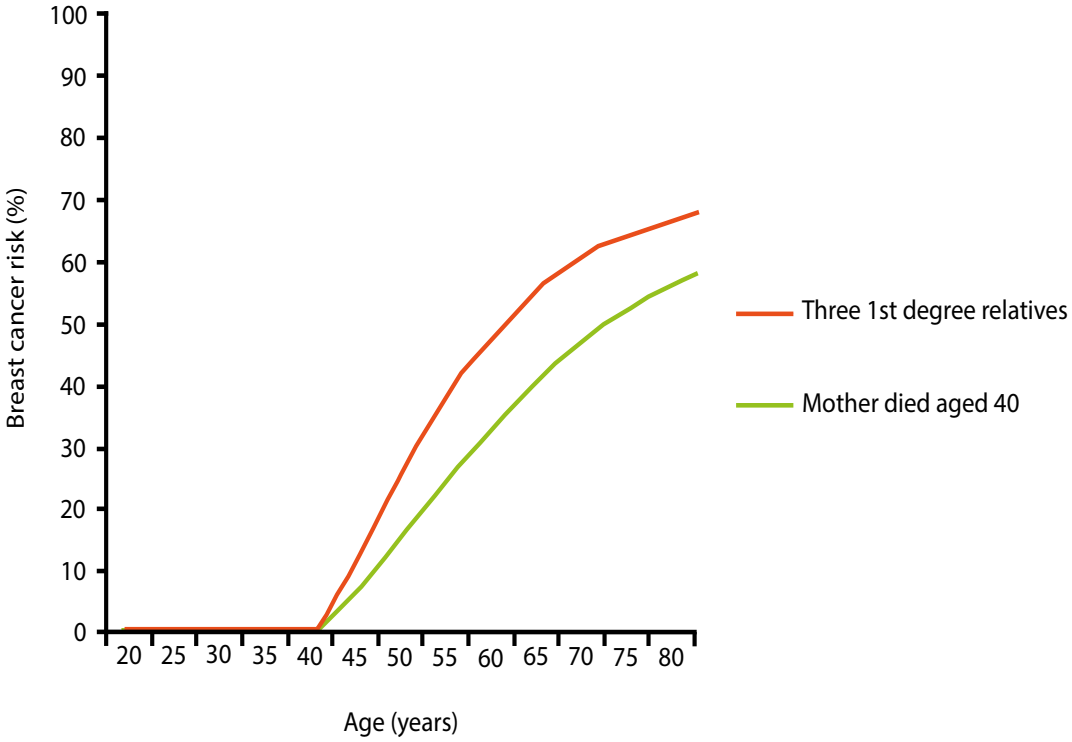


Figure 2b: Breast cancer risks associated with the same *BRCA1* mutation arising in women tested at the same age who have different family histories (based on BOADICEA risk calculations 4 and 10)



Published estimates of penetrance range from around 57% to 71% at age 70 years for a *BRCA1* mutation and 45% to 87.5% for *BRCA2* – table 3. The variation reflects both the effect of risk modifiers and the ascertainment of mutation carriers on which the estimates are based. In general, estimates are higher in studies based on multi-case families than those based on unselected breast cancer cases, in part due to ascertainment bias [40].

For a woman in the general population, these figures are likely to be over-estimates of the true absolute breast cancer risk associated with carrying a *BRCA* mutation: in unselected or 'average risk' populations, the most appropriate penetrance estimate would be the average penetrance across all possible modifiers of breast cancer risk. This was estimated as 59% for *BRCA1* mutation carriers born after 1950 and 51% for *BRCA2* mutation carriers in a study that used complex segregation analyses using data from both multi-case families and breast cancer case series to model explicitly the contribution of risk modifiers [38].

Table 3: Published estimates of penetrance of *BRCA1* and *BRCA2* mutations

Approach	Penetrance of *BRCA1* mutation (%)	Penetrance of *BRCA2* mutation (%)	Reference
Modelling of breast cancer incidence in average risk women with population carrier frequency of 1 in 397	68.6 (47.7-83.9) at age 75 years	No data at age 75 years	Nelson 2005 [36]
Meta-analysis of mean cumulative cancer risk for mutation carriers in 10 studies from high risk families and unselected patients	57 (47-66) at age 70 years	49 (40-57) at age 70 years	Chen 2007 [41]
Meta-analysis of mean cumulative cancer risk in 22 studies involving 8139 index cancer patients unselected for family history	65 (51-75) by age 70 years	45 (33-54) by age 70 years	Antoniou 2003 [40]
Complex segregation analysis based on the 22 population-based studies above and additional family data that modelled penetrance across all risk modifiers	59 by age 70 years for those born after 1950	51 by age 70 years for those born after 1950	Antoniou 2008 [38]
Cumulative incidence of cancer derived from a prospective series of 988 *BRCA1/2* carriers without disease at baseline in the UK	60 (44-75) by age 70 years	55 (41-70) by age 70 years	Mavaddat 2013 [42]
Penetrance estimated in 237 families with at least four breast cancers in the Breast Cancer Linkage Consortium	71 (53-82) by age 70 years	84 (43-95) by age 70 years	Ford 1998 [43]
Penetrance estimated in 1188 female members of 185 *BRCA1/2* families in Netherlands	71.4 (67.2-82.4) by age 70 years	87.5 (82.4-92.6) by age 70 years	Van Der Kolk 2010 [44]

2.3 *PALB2*

Location, function and role in cancer pathogenesis

PALB2 (Partner and Localiser of *BRCA2*) is located on chromosome 16p12. It was first identified as a key protein that interacts with *BRCA2* to assist with DNA repair and related genome caretaker functions [45]. Loss of function mutations in both *PALB2* alleles cause Fanconi's anaemia, whereas monoallelic loss of function mutations are associated with an increased risk of breast and pancreatic cancer. Most loss of function mutations are frameshift, nonsense or splice site mutations with a very few due to genomic deletions [46].

Population prevalence of PALB2 mutations

Population frequencies of *PALB2* mutations vary depending on ethnicity and geography, with several founder mutations reported in areas such as Central Europe and Scandinavia. Very few studies examine prevalence of *PALB2* mutations in unselected populations.

Penetrance of PALB2 mutations

Two studies using modified segregation analysis estimated the absolute risk of breast cancer by age 70 in women with *PALB2* mutations as 35-40% [54]; [56]. Absolute risk was calculated by multiplying measured genetic relative risk (i.e. risk in the carrier compared to risk in the average non carrier woman in population) by polygenic risk estimated from family history (which centres on the average population risk for that age). Relative risk estimates quoted for heterozygous mutation carriers vary depending on the study population but appear to be at least five or six times that of the general population. A study showing only a 2.3 fold increase in breast cancer risk for *PALB2* mutation carriers [51] was criticised for failing to take into account that women with a strong family history of breast cancer will be at the highest end of the polygenic risk scale, even before considering the additional effects of *PALB2* [53].

Table 4: Prevalence of *PALB2* mutations in non-cancer populations

Population	Proportion of (non-cancer) population who are *PALB2* mutation carriers	Allele frequency	Reference
Czech			
Non cancer controls in a study of familial breast cancer	1 in 1,226 (c.509_510delGA)	0.0004	Janatova 2013 [47]
Poland			
Healthy blood donor controls in a study of breast and ovarian cancer	1 in 1,310 (c.509_510delGA)	0.00038	Dansonka-Mieszkowska 2010 [48]
Finland			
Healthy blood donor controls in a study of familial breast cancer	6 in 2,501 (1592delT)	0.0012	Erkko 2007 [49]
Healthy population controls in a study of familial and sporadic breast cancer	2 in 1,079 (1592delT)	0.0009	Heikkinen 2009 [50]
UK			
Healthy controls in a study of familial breast cancer	0 in 1,084 (five different mutations detected in cases)	0	Rahman 2006 [51]
Canada			
Newborn healthy French-Canadian controls	0 in 6,440 (most common mutation in cases = 2323C>T)	0	Foulkes 2007 [52]

Table 5: Penetrance of *PALB2* mutations in non-cancer populations

Approach	Penetrance *PALB2* mutation(%)	Effect estimate	Reference
Risk of breast cancer analysed in 362 members of 154 families with *PALB2* mutations. Modified segregation-analysis approach used to allow for the effects of *PALB2* genotype and residual familial aggregation (worldwide)	35 (95% CI, 26 to 46) by age 70 years if family history not taken into account Note this estimate increased with a strong family history of breast cancer up to 58 (95% CI, 50 to 66) if mother and sister had breast cancer by age 50 years	RR 9.07 (95% CI, 5.72 to 14.39)	Antoniou 2014 [54]
Case control study of 559 cases with contralateral breast cancer and 565 controls with unilateral breast cancer from the WECARE study (cancer registry Denmark and some US states)	-	RR 5.3 (95% CI, 1.8 to 13.2)α	Tischkowitz 2012 [55]
Modified segregation analysis using 1718 breast cancer cases unselected for family history or age at diagnosis (Finland)	40% (95% CI, 17 to 77) by age 70 years	HR 6.1 (95% CI, 2.2 to 17.20) NB HR decreased with age	Erkko 2008 [56]
Case control study of breast cancer families and unselected breast cancer cases (Finland)	-	OR 11.3 (95% CI, 1.8 to 57.8) in high risk families OR 3.94 (95% CI, 1.5 to 12.1) in unselected cases	Erkko 2007 [49]
Segregation analysis using info from 1,084 non-cancer controls and full pedigrees of affected individuals (923 UK patients with familial breast cancer)	-	RR 2.3 (95% CI, 1.4 to 3.9)	Rahman 2006 [51]
Australian Breast Cancer Family Registry of families with early onset breast cancer	49% (15 to 93) by age 50 91% (44 to 100) by age 70	HR 30.1 (95% CI, 7.5 to 120)	Southey 2010 [57]

α Breast cancer risk estimated by comparing breast cancer incidence in first degree female relatives of carrier cases with incidence in first degree relatives of non-carrier contralateral breast cancer cases using 'conventional actuarial techniques and then by transforming the estimate'
RR = relative risk, HR = hazard ratio, OR = odds ratio

2.4 *TP53*

Location, function and role in cancer pathogenesis

TP53 is located at 17p13.1. It encodes a tumour suppressor protein that plays a key role in regulating the cell cycle in response to cellular stresses e.g. to induce cell cycle arrest, apoptosis, senescence, DNA repair or metabolic changes [58]. Germline mutations in TP53 are associated with predisposition to a spectrum of cancers including those linked to the mutiple neoplasm syndrome Li Fraumeni syndrome. Four tumours account for 77% of all Li Fraumeni-associated cancers; these are breast cancer, sarcomas, brain tumours and adrenocortical carcinoma [59]. Germline *TP53* mutations have also been detected in patients with breast cancer who do not meet the criteria for Li Fraumeni sydrome.

Population prevalence of TP53 mutations

Studies of prevalence of *TP53* mutations tend to be carried out in families with Li Fraumeni syndrome so ascertaining population prevalence of mutations is challenging and estimates from the few existing studies are highly variable. A Brazilian founder mutation in *TP53* has been reported (R337H) that is present in the majority of childhood adrenocortical carcinoma but its role in breast carcinogenesis is unclear.

Table 6: Prevalence of *TP53* mutations in non-cancer populations

Population	Proportion of (non-cancer) population who are *TP53* mutation carrier	Allele frequency	Reference
California, USA population prevalence calculated from data on individuals sent for genetic analysis of *TP53* in a clinical genetics lab and early onset breast cancer patients	1 in 20,000	0.000025	Gonzalez 2009 [60]
UK population prevalence calculated from data on early onset breast cancer patients	1 in 5,000	0.0001	Lalloo 2003 [61]
Newborns in Southern Brazilians offered screening for germline *TP53* mutation R337H	461 in 171,649	0.0013	Custodio 2013 [62]

Penetrance of TP53 mutations

Carriers of germline *TP53* mutations from Li Fraumeni families have a very high risk of developing cancers including breast cancer at a relatively young age. The lifetime risk of any cancer for female carriers of these *TP53* mutations approaches 100% , with the majority of cases due to breast cancer. Data are not readily available on absolute breast cancer risk in people with germline *TP53* mutations who are not from Li Fraumeni families.

Table 7: Penetrance of *TP53* mutations in non-cancer populations

Approach	Penetrance of *TP53* mutation in breast cancer (%)	Effect estimate	Reference
Review (but basis for statement unclear)	49% in women by age 60 years	-	Maschiari 2012 [63]
Germline *TP53* mutations were sought in 268 index cases of childhood cancer with a relative affected by early onset cancer or affected by multiple cancers. Maximum likelihood methods correcting for ascertainment bias were used to estimate lifetime cancer risks	84% in women by age 45 years 100% in women by age 85 years For any LFS-type cancer, although breast cancer formed the majority of cases	-	Chompret 2000 [64]
Cancer incidence data collected on 107 index cases who survived childhood soft tissue sarcomas and 3,257 of their relatives and compared in those with and without *TP53* mutations	-	SIR for breast cancer 105.1 (55.9–179.8)	Hwang 2003 [65]

SIR = standardised incidence ratio

2.5 Test performance of WGS for detecting inherited breast cancer variants

Accurate knowledge of the absolute cancer risk associated with a mutation in a particular population will allow gene variants conferring risks above a certain threshold to be fed back to clinicians and patients to enable appropriate clinical action. First, however, a test for inherited breast cancer gene variants aims to identify and classify correctly any variants detected. To evaluate the performance of WGS for detecting pathogenic mutations in genes such as *BRCA1* and *BRCA2* – used here as an example because they are relatively well-described – the following areas should be considered.

2.6 Analytical validity

Analytical validity refers to laboratory performance of an assay, i.e. how accurately and reliably the assay measures the genotype of interest, here this refers to testing of all pathogenic mutations in *BRCA*1 and *BRCA2*. To optimise clinical diagnostic value, WGS should detect clinically actionable variants of all classes (such as single nucleotide variants, indels, translocations and large deletions) with very high sensitivity and specificity [66]. Major components involved in WGS that may affect analytical validity include:

- **Generating enough high quality raw DNA sequence** [2]. The number and quality of fragments of DNA available will influence the depth of coverage (number of times the genome is sequenced in any one individual, also known as 'reads') as well as horizontal coverage across the genome. A greater number of reads increases statistical confidence that a given DNA base in the sequence has been inferred correctly. Although there has been a widespread view that a 30x average depth of coverage is necessary to ensure adequate sensitivity and specificity for clinical use, recent evidence suggests that a mean read depth of 14x may be sufficient [67]
- **Accurately assembling the patient's genome and identifying variants.** When a genome sequence is assembled from multiple DNA fragments, some types of disease-causing genomic variants such as large deletions and copy number variants cannot be uniquely positioned on the reference genome map, and therefore cannot be effectively identified by whole genome sequencing [2]. The choice of reference genome will also impact on variants identified
- **Filtering and prioritising potentially pathogenic variants.** Knowledge of all potentially pathogenic variants is incomplete so rigour is needed to ensure that non-functional or normal variants are not interpreted as being causal [2]
- **Interpreting, validating and reporting clinically significant variants.** Manual review of the final list of potentially pathogenic variants by experienced clinicians and scientists is essential to judge whether there is sufficient evidence to support pathogenicity and thereby reduce the risk of reporting false positives [2]. At present, for any potential pathogenic variants identified by WGS, accuracy should be validated by comparing with sequences generated through different methods such as Sanger sequencing, though as data accumulate this is likely to change

Laboratory reporting of sequence variants

Interpreting sequence variants as high or low risk presents a significant challenge. In the US, the ACMG recommends adopting a six-category approach to clinical reporting of genetic sequence variation [68]:

- Category 1: Sequence variation is previously reported and is a recognised cause of the disorder (mutation is likely pathogenic)
- Category 2: Sequence variation is previously unreported and is of the type which is expected to cause the disorder (mutation is likely pathogenic)
- Category 3: Sequence variation is previously unreported and is of the type which may or may not be causative of the disorder (mutation is VUS)
- Category 4: Sequence variation is previously unreported and is probably not causative of disease (mutation is likely benign)
- Category 5: Sequence variation is previously reported and is a recognised neutral variant (mutation is likely benign)
- Category 6: Sequence variation is not known or expected to be causative of disease, but is found to be associated with a clinical presentation (mutation significance is unclear)

In the UK, the Association for Clinical Genetic Sciences (ACGS) recommends a five-category approach as outlined below [69] (with classes numbered in reverse order to ACMG categories).

Table 8: ACGS recommendations for classification of sequence variants

Class	Description	Wording to include in reports
1	Clearly not pathogenic	Not pathogenic "Common" polymorphism and therefore not reported
2	Unlikely to be pathogenic	Unlikely to be pathogenic Diagnosis not confirmed molecularly
3	Unknown significance (VUS)	Uncertain pathogenicity Does not confirm or exclude diagnosis
4	Likely to be pathogenic	Likely to be pathogenic Consistent with the diagnosis
5	Clearly pathogenic	Predicted to be pathogenic This result confirms the diagnosis

Plon *et al* [70] propose adapting the ACGS categories to attach a probability to the likelihood that a variant is pathogenic is shown in the table below.

Table 9: Proposed classification system for sequence variants identified by genetic testing

Class	Description	Probability of being pathogenic
1	Not pathogenic or of no clinical significance	<0.001
2	Likely pathogenic or of little clinical significance	0.001-0.049
3	Uncertain	0.05-0.949
4	Likely pathogenic	0.95-0.99
5	Definitely pathogenic	>0.99

Regardless of the classification system used, categories of sequence variants can generally be grouped into three 'bins': variants that are likely to be pathogenic (positives); variants that are likely to be benign (negatives) and variants of uncertain significance (VUS). This final category remains a source of disagreement between experts [71] and differing practice between laboratories. Although typically around 10-15% of variants detected from high risk families are VUS, this figure may be very different for low risk populations where the majority of variants detected are likely to be VUS. The effect of the size of the VUS pool on WGS for *BRCA1* and *BRCA2* will be explored below.

Mutation spectrum in BRCA1 *and* BRCA2

1,734 discrete mutations in *BRCA1*, not all of them pathogenic, were reported in the Leiden Open Variation Database as of 27 October 2014 [72], of which the great majority (over 90%) were due to substitutions. These variants are summarised in the table below.

Table 10a: Number and type of *BRCA1* mutations reported in the Leiden Open Variation Database

Type of Variant	Number (%)	Location			
		5′ Start	coding	intron	3′ stop
Substitutions	1616 (93.2)	1	1476	139	0
Deletions	41 (2.4)	0	27	14	0
Duplications	11 (0.6)	0	9	2	0
Insertions	4 (0.2)	0	1	3	0
Insertion / Deletions	1 (0.06)	0	0	1	0
Complex	26 (1.5)	-	-	-	-
Unknown	35 (2.0)	-	-	-	-
Totals	1734 (100.0)	1	1513	159	0

Similarly there were 1,020 variants reported in *BRCA2*, of which 90.6% were substitutions.

Table 10b: Number and type of *BRCA2* mutations reported in the Leiden Open Variation Database

Type of Variant	Number (%)	Location			
		5′ Start	Coding	Intron	3′ Stop
Substitutions	924 (90.6)	5	826	93	0
Deletions	39 (3.8)	0	28	11	0
Duplications	2 (0.2)	0	1	1	0
Insertions	3 (0.3)	0	1	2	0
Insertion/Deletions	1 (0.1)	0	0	1	0
Complex	9 (0.9)	-	-	-	-
Unknown	43 (4.2)	-	-	-	-
Totals	1020 (100.0)	5	856	107	0

A recent population series of 2,222 ovarian cancer cases and 1,528 controls characterised the spectrum of pathogenic mutations further: 85 *BRCA1* mutations were detected, comprising 51 frameshift indels and 34 missense, nonsense or splice site mutations; there were also 98 *BRCA2* mutations, made up of 75 frameshift indels and 23 missense, nonsense or splice site mutations [73]. In addition around 10-12% are due to large rearrangements / duplications [31]. We therefore assumed that relative proportions of pathogenic *BRCA1* mutations were 54% small indels, 36% SNVs and 10% CNVs; for pathogenic *BRCA2* mutations the corresponding figures were 69% small indels, 21% SNVs and 10% CNVs.

WGS is likely to have good potential to detect the 90% of *BRCA1* and *BRCA2* mutations that are due to small sequence variants (in particular substitutions) but may perform less well for large rearrangements and duplications, although this is a rapidly developing area. This falls short of the 95% sensitivity that is aimed for with current genetic diagnostic tests [74]. In contrast the majority of VUS are missense variants, so most should be detected by WGS.

Laboratory parameters

Very few laboratories worldwide currently offer WGS for clinical purposes so methods tend not to be standardised, although general laboratory standards for next generation sequencing including WGS have been published by the ACGS [74] and ACMG [75]. Assay performance therefore varies by laboratory according to the method and machine used for generating the sequence and programmes used to display output (align, call and annotate variants). There is no published information validating WGS test performance specifically for *BRCA1* and *BRCA2* variants against Sanger sequencing.

Illumina example

Illumina has a TruGenome Clinical Sequencing Service that offers WGS for 1) finding genetic causes of undiagnosed disease, 2) investigating carrier status and genetic predisposition to adult-onset conditions for healthy people, 3) offering technical sequence data to labs and physicians able to make their own clinical interpretation. The second of these is known as the TruGenome Predisposition Screen and covers a set of 1,232 conditions caused by variants in well-characterised single genes [76] including *BRCA1* and *BRCA2*. Illumina publish information on the proportion of sequence variants and deletions identified by their WGS algorithms (>97% and >80% respectively). While horizontal coverage across the genome is typically 95% for WGS , Illumina report 99.41% coverage in the region of *BRCA1* and 99.97% for *BRCA2*. Depth of coverage is to an average of 30x with a minimum depth of 10x.

Test literature states that the test does not detect insertion / deletion events >7bp, translocations, trinucleotide repeats or copy number variants. This is, however, another area of rapid development so it is likely that some of these variants including large rearrangements will be detected in future with development of appropriate algorithms.

2.7 Methods for test performance calculations

We used information from scientific literature on *BRCA1* and *BRCA2* mutation prevalence in unselected populations combined with Illumina published estimates of analytical validity to estimate the sensitivity, specificity, positive and negative predictive values of WGS for detecting pathogenic variants in an unselected population of 100,000 UK women. Details were published in a paper in Hereditary Cancer in Clinical Practice in June 2016 [77]. Model input parameters are shown in table 11.

Table 11: Input parameters for calculations of WGS test performance for detecting pathogenic *BRCA1* and *BRCA2* mutations

Input Parameters	Values for *BRCA1*	Values for *BRCA2*
Prevalence of pathogenic mutations in unselected population	0.0012	0.002
Theoretical population size	100000	100000
Proportion of pathogenic mutations that are small indels	0.54	0.69
Proportion of pathogenic mutations that are SNVs (nonsense, pathogenic missense, splice site)	0.36	0.21
Proportion of pathogenic mutations that are CNVs	0.1	0.1
Number VUS for every pathogenic variant	variable	variable
Gene coverage using WGS	0.9941	0.9997
Sensitivity of WGS for small indels	0.8	0.8
Sensitivity of WGS for SNVs	0.97	0.97
Sensitivity of WGS for CNVs	0	0
Specificity of WGS for indels	1	1
Specificity of WGS for SNVs	1	1
Specificity of WGS for CNVs	1	1

Analytic validity calculations shown below for *BRCA1* were repeated for *BRCA2*. We assumed that variants of uncertain significance (VUS) were not reported back to patients in line with common practice [78].

- True positives = prevalence of pathogenic *BRCA1* mutations x population size x (proportion of *BRCA1* mutations that are small indels x sensitivity of WGS for detecting small indels + proportion of *BRCA1* mutations that are SNVs x sensitivity of WGS for detecting SNVs + proportion of *BRCA1* mutations that are CNVs x sensitivity of WGS for detecting CNVs) x horizontal gene coverage of WGS for *BRCA1*
- False negatives = prevalence of pathogenic *BRCA1* mutations x population size – true positives
- False positives = in main model assumed to be 0 as mutations detected on WGS would be confirmed by a different method e.g. Sanger sequencing. In sensitivity analysis we modelled the effect of an extra 0-5% of false positive results in addition to true positive results
- True negatives (including VUS) = population size - prevalence x population size - false positives

Test performance measures of WGS for detecting pathogenic mutations were calculated as follows for *BRCA1* then repeated for *BRCA2*:

- Analytical sensitivity = True positives detected / Total with a true pathogenic mutation
  i.e. prevalence x population size
- Analytical specificity = True negatives detected / Total without a pathogenic mutation
  i.e. population size - prevalence x population size
- Analytical positive predictive value = True positives detected / Total with a variant on testing
  i.e. True positives + False positives
- Analytical negative predictive value = True negatives detected / Total without a variant on testing
  i.e. True negatives + False negatives

Sensitivity analysis

We also performed a sensitivity analysis to investigate the effect of varying model input parameters. The model was rerun 100 000 times with model input parameters being randomly selected from defined likely distributions using Stata's random number generator function. The proportion of pathogenic mutations due to CNVs was assumed to be fixed at 0.1, but the proportion of small indels and SNVs varied according to an underlying normal distribution. Sensitivity of WGS for detecting CNVs was fixed at 0, based on current test performance literature, but sensitivity for detecting SNVs and small indels was selected from an underlying gamma distribution. We also assumed that false positives would occur at a rate of less than 10% of the number of true positives, but heavily skewed towards 0. Calculations for true positives, false negatives and true negatives remained the same as for the main model. Underlying distributions for model input parameters are shown in figure 3.

2.8 Results of test performance calculations

In our main model, testing would successfully detect 93 women out of 120 with pathogenic *BRCA1* variants and 151 women out of 200 with pathogenic *BRCA2* variants in a hypothetical UK population of 100 000 women. Sensitivity analysis gave similar results – table 12.

This resulted in an analytic sensitivity of 77.5% for *BRCA1* variants and 75.5% for *BRCA2*. Specificity was 100%, with 100% positive predictive value (PPV) and 99.9% negative predictive values for both genes. It is clear that the sensitivity of WGS for detecting pathogenic *BRCA1* and *BRCA2* mutations is sub-optimal. The number of false positives detected by WGS and reported back to patients and clinicians, however, is likely to be negligible given that any potentially pathogenic mutation would be verified using an alternate technology.

In a model situation where VUS are considered to be negative and are not reported back to patients, the PPV of a positive result on WGS approaches 100%. The PPV will however vary considerably depending upon the size of the VUS pool and how VUS are reported. This is explored in tables 13 and 14 in the context of *BRCA1* which show the effects of having one VUS for each pathogenic mutation and 10 VUS for each pathogenic mutation respectively. The true situation is unknown.

Figure 3: Distribution of model input parameters used for sensitivity analysis using example of *BRCA1*

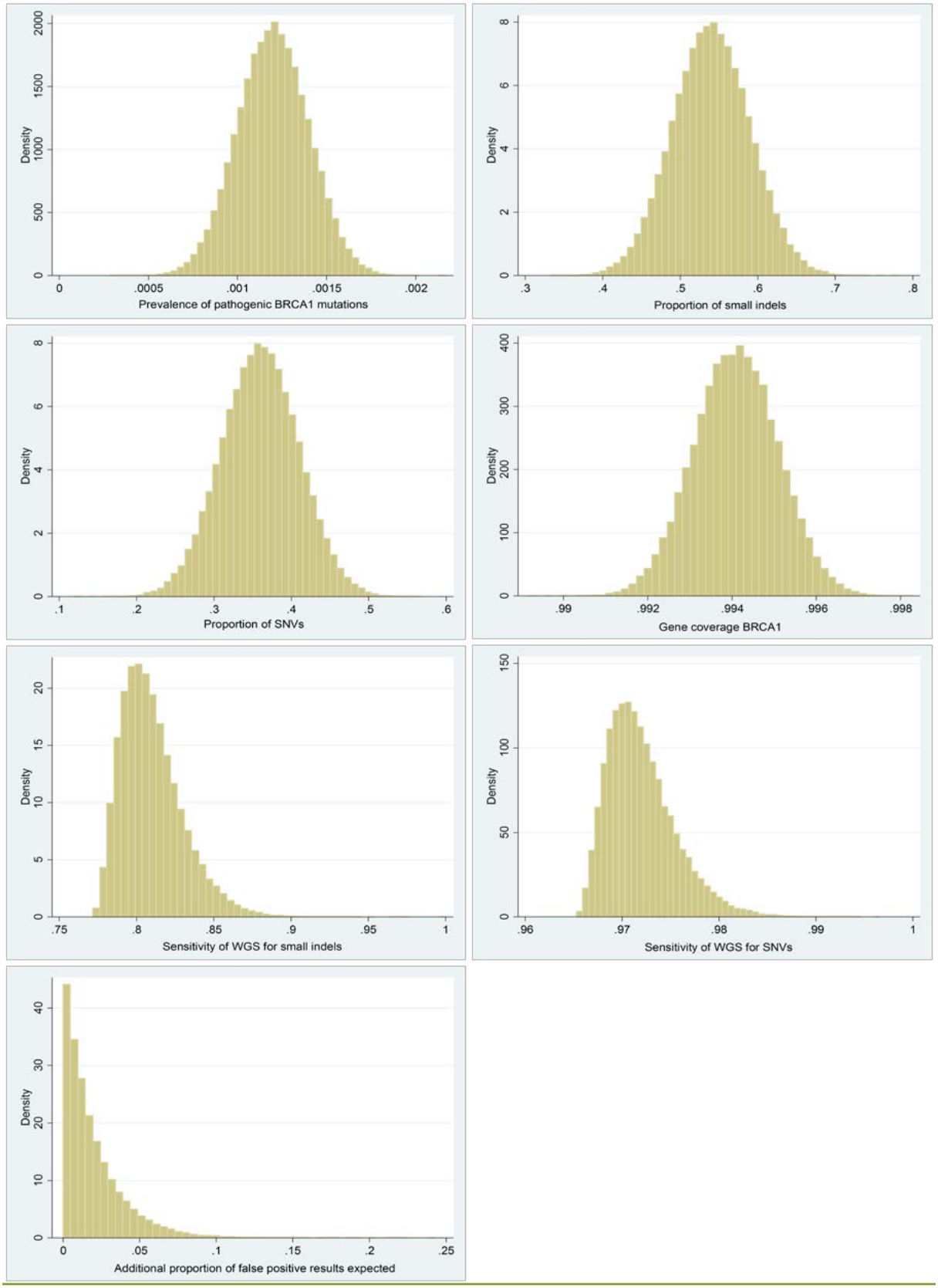


Table 12a: Numbers of pathogenic *BRCA1* mutations detected in an unselected population of 100,000 UK women using WGS

BRCA1 main model	Has gene variant?			
Variant detected by WGS?	Yes	No	Total	
Yes	**93**	**0**	**93**	
No	**27**	**99,880**	**99,907**	
Total	**120**	**99,880**	**100,000**	
BRCA1 Sensitivity analysis	Mean (SD)	Median (IQR)	Min	Max
True Positives	94 (15.8)	94 (83-105)	26	175
False Positives	2 (1.9)	1 (0-3)	0	26
False Negatives	26 (4.7)	26 (23-29)	7	48
True Negatives	99,878 (20.4)	99,878 (99,864 - 99,891)	99,777	99,967

Table 12b: Numbers of pathogenic *BRCA2* mutations detected in an unselected population of 100,000 UK women using WGS

BRCA2 main model	Has gene variant?			
Variant detected by WGS?	Yes	No	Total	
Yes	**151**	**0**	**151**	
No	**49**	**99,880**	**99,849**	
Total	**200**	**99,880**	**100,000**	
BRCA2 Sensitivity analysis	Mean (SD)	Median (IQR)	Min	Max
True Positives	153 (23.0)	153 (137-168)	49	245
False Positives	3 (3.1)	2 (1-4)	0	37
False Negatives	48 (7.8)	47 (42-53)	16	88
True Negatives	99,797 (30.5)	99,796 (99,776 - 99,817)	99,672	99,934

Table 13: Test performance of WGS for detecting pathogenic *BRCA1* mutations in a population of 100,000 women, assuming that one missense VUS is reported for each pathogenic *BRCA1* mutation

BRCA1	Has gene variant?			
Gene variant detected by WGS?	Yes	VUS	No	Total
Yes	93	116	0	
No	27	4	99,760	
Total	120	120	96,760	100,000

Table 14: Test performance of WGS for detecting pathogenic *BRCA1* mutations in a population of 100,000 women, assuming that 10 missense VUS are reported for each pathogenic *BRCA1* mutation

BRCA1	Has gene variant?			
Gene variant detected by WGS?	Yes	VUS	No	Total
Yes	93	1,157	0	1,250
No	27	43	98,680	98,750
Total	120	1,200	96,680	100,000

Although the presence or absence of VUS does not affect the analytic sensitivity or specificity of WGS for detecting pathogenic mutations, there is potential for VUS to result in costly extra laboratory work depending on the scale of attempts made to classify the variant. While most VUS are likely to be non-pathogenic, this is especially true for populations at low prior risk of pathogenic *BRCA1* and *BRCA2* mutations.

Other considerations affecting WGS test performance for detecting inherited breast cancer variants

Laboratory variation in sequencing, interpretation and reporting

The issues of differing coverage and concordance between sequencing platforms used for WGS have been highlighted in the literature [2]; [79]. Consistency in the interpretation and classification of variants is essential, especially if WGS is to be done at scale across different laboratories. To reduce the risk of false positives, the presence of any pathogenic mutations should be confirmed using a different technology such as Sanger sequencing before reporting results back to a patient. Specific strategies are needed for the interpretetation of VUS, as differences in the way VUS are dealt with impacts substantially on a test's PPV.

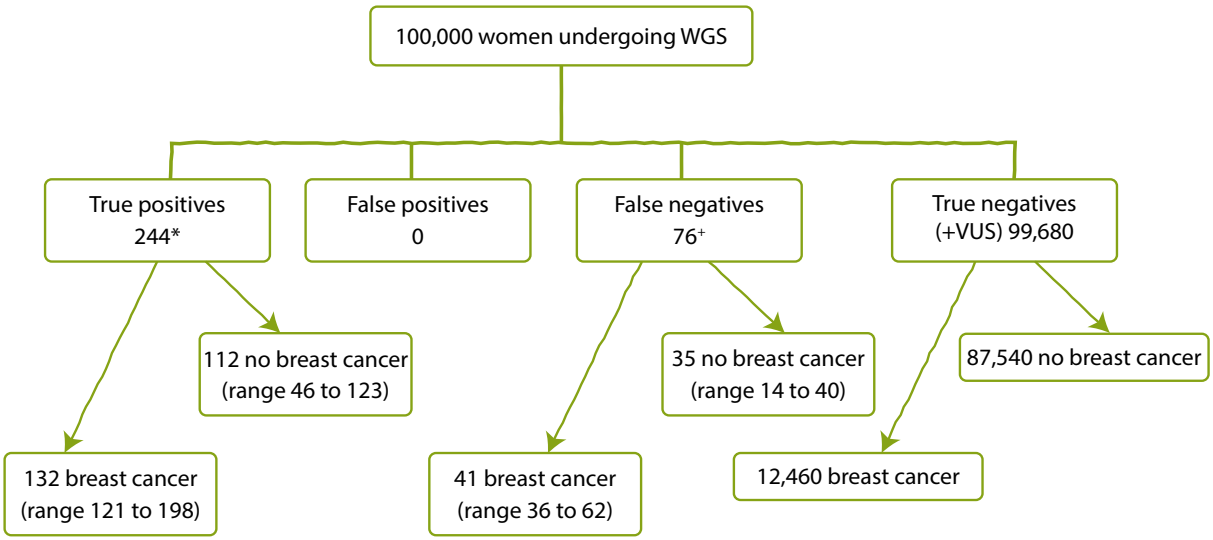
Changes over time

Developments in sequencing technology, software and algorithms used to report variants mean that the test performance figures outlined above are likely to change over time. The potential for whole genome data to be re-analysed or re-interpreted at a later date means that an individual patient's results may also change over time. For simplicity we have modelled test performance based on testing individuals at one time point only.

2.9 Clinical validity

Clinical validity here refers to the performance of WGS for predicting future risk of breast cancer. This depends both on aspects of test performance such as sensitivity, specificity, positive and negative predictive values as well as on the association between genotype and disease [24] (penetrance) in this theoretical population of UK women undergoing WGS.

Figure 4: Flowchart of expected breast cancer incidence in unselected women undergoing WGS



*Comprises 93 women with *BRCA1* mutations and 151 with *BRCA2* mutations
+Comprises 27 women with *BRCA1* mutations and 49 with *BRCA2* mutations

In our model, WGS would correctly identify 132 women (range 121-198) with a pathogenic mutation who then developed breast cancer. We estimated that breast cancer would also occur in 41 women (range 36-62) incorrectly identified with no pathogenic *BRCA1* or *BRCA2* mutations (false negatives) and in 12,460 true negatives. Estimates were based on a 12.5% background lifetime risk of female breast cancer in the UK [26] as well as the assumption that 59% [38] (range 57% to 71%) of women with pathogenic *BRCA1* mutations and 51% [38] (range 45% to 87.5%) of women with pathogenic *BRCA2* mutations would develop breast cancer by age 70 years.

2.10 Clinical utility

Before considering possible health outcomes associated with testing women for inherited breast cancer variants using WGS, underlying assumptions should be clarified. First, we assume that a woman has received appropriate genetic counselling prior to the test, that she is aware of the possibility of secondary findings of high risk breast cancer variants and has given consent for WGS and return of secondary findings. Second, we assume that any potentially pathogenic mutation detected by WGS is verified using a different method. Third, we do not consider the situation in which test results change after re-analysis or re-interpretation of data due to new genetic technology or understanding. As this theoretical population is undergoing WGS anyway, added costs of sequencing *BRCA1* and *BRCA2* are assumed to be negligible, although sequence interpretation may incur an additional cost.

Potential health outcomes

Scenario one: Analytic true positive women = 244/100 000

These women have a high lifetime risk of breast cancer that can potentially be mitigated by intervention so WGS has potential to produce a positive health outcome.

Scenario two: Analytic false positive women = 0/100 000

These women have a low lifetime risk of breast cancer but may still be offered intervention, which may result in negative health outcomes.

Scenario three: Analytic false negative women = 76/100 000

Although these women have a high lifetime risk of breast cancer, WGS can result in no positive health outcomes for them. There are also unlikely to be any major negative consequences.

Scenario four: Analytic true negatives (includes VUS) = 99 680/100 000

There are no direct positive or negative health outcomes for this group. For those with VUS, lifetime breast cancer risk is uncertain so health outcomes cannot be assessed.

Description of health outcomes

Scenario one health outcomes

For the 132 'true positive' women predicted to develop breast cancer, potential positive health outcomes include:

- Reduction in breast cancer incidence and mortality through interventions offered
- Identification of family members at high risk with subsequent reductions in their breast cancer incidence and mortality

Potential harms or negative health outcomes for this group include:

- Psychological distress associated with being labelled as 'high risk'
- Psychological distress of repeated screening
- Physical harms from chemopreventive treatments e.g. side effects, thromboembolism
- Physical and / or psychological harms from prophylactic surgery e.g. pain, scarring, effects on body image

Of note, the 112 other analytic true positive women would have the same preventive options but receive no benefit.

Scenario two health outcomes

For 'false positive' women there are no potential positive health outcomes. All negative health outcomes identified for true positive women would apply, although numbers of women affected are likely to be extremely small.

Scenario three health outcomes

For 'false negative' women, there are no potential positive health outcomes.

Potential negative health outcomes for this group include:

- False reassurance associated with receiving a negative result
- Missed opportunities to prevent breast cancer

Scenario four health outcomes

For 'true negative' women or those with VUS, there are no direct positive or negative health outcomes. There is a risk that without careful consent and clear clinical feedback some may be falsely reassured that they are not at risk of breast cancer which might affect subsequent health behaviour.

Description of other outcomes

Although there is a wide range of potential psychological, ethical, legal and social implications of using WGS to test for inherited breast cancer variants in women at low prior risk of breast cancer, full consideration of these is outside the scope of this review.

Current clinical practice for women with BRCA1 *and* BRCA2 *mutations*

Evidence-based guidance from NICE for women with identified *BRCA1* and *BRCA2* mutations (scenarios 1 and 2 above) and no personal history of breast cancer recommends the following actions [18]:

- Enhanced surveillance for breast cancer with the screening modality dependent on age (table 16)
- Lifestyle advice on smoking, weight, physical activity, breast awareness and breastfeeding in line with general health advice offered to all women
- Discussion of personal risks and benefits of OCP (if relevant)
- Discussion of personal risks and benefits of HRT (if relevant)
- Consideration of chemoprevention using tamoxifen or raloxifene
- Consideration of bilateral risk-reducing mastectomy for high risk women

Table 16: Surveillance recommended for *BRCA* positive women with no personal cancer history

Age Group	Surveillance recommended
20-29	Do not offer mammography and do not offer MRI
30-39	Annual MRI and conisder annual mammography
40-49	Annual mammography and annual MRI
50-59	Annual mammography Do not offer MRI unless dense breast pattern
60-69	Annual mammography Do not offer MRI unless dense breast pattern
70+	Mammography as part of the population screening programme

Assessment of interventions to prevent breast cancer in women with BRCA1 *and* BRCA2 *mutations*

Chemoprevention positive effects

In an extensive evidence review conducted by the US Preventive Task Force in 2005 [36], updated in 2014 [80], six placebo-controlled trials (four tamoxifen, two raloxifene) and one head-to-head trial showed that tamoxifen and raloxifene reduced the incidence of invasive breast cancer by 30% to 68% compared with placebo. The number of *BRCA* mutation carriers needed to treat with chemoprevention to prevent one case of breast cancer was 3.9 (95% C.I. 2.6-9.1) [36]. It should be noted that positive health outcomes can only be realised if a woman takes up the offer of preventive treatment. The proportion of women choosing chemoprevention was estimated at between 5% and 50% [36].

Chemoprevention negative effects

The incidence of side effects associated with chemoprevention was estimated as 0.096% per year for thrombo-embolic events (95% C.I 0.036-0.156) and 0.036% per year for endometrial cancer risk (95% C.I 0.00177-0.0709) [36]. The numbers of *BRCA* mutation carriers needed to treat to cause one thrombo-embolic event and one case of endometrial cancer per year were 1,042 and 2,686 respectively [36].

Surgery positive effects

The US Preventive Task Force found that among high risk women and mutations carriers, risk-reducing mastectomy decreased breast cancer incidence by 85% to 100% compared with women without surgery and salpingo-oophorectomy reduced breast cancer by 37%-100% according to four studies of mastectomy and three of oophorectomy or salpingo-oophorectomy [80]. The number of *BRCA* mutation carriers needed to treat with mastectomy to prevent one case of breast cancer was 1.6 (95% C.I. 1.3-2.4) [36]. The corresponding number needed to treat with oophorectomy to prevent one case of breast cancer was 2.2 (95% C.I. 1.5-148) [36]. Significantly reduced anxiety after surgery was seen in a study of 90 high risk women undergoing risk-reducing mastectomy. Between 5% and 20% of women chose mastectomy compared to 25% to 75% who accepted oophorectomy [36].

Surgery negative effects

Potential negative effects of surgery include surgical complications, long-term physical and psychological effects. In the earlier review overall complication risk was quoted as 21% for risk-reducing mastectomy and 5% for oophorectomy [36].

3. Conclusions

At present there are many uncertainties that will impact on the clinical utility of actively seeking secondary findings using WGS, even for relatively well-characterised genes such as *BRCA1* and *BRCA2*. The process demonstrated in this paper when applied to other gene-disease combinations is likely to highlight further gaps in knowledge and understanding around clinical validity and utility, which should be considered before offering routine testing of genes for secondary purposes in unselected populations.

This paper demonstrates the type of process that should be undertaken when considering likely test and health outcomes that may arise when using WGS in unselected populations, using the example of inherited breast cancer variants. The aim of detecting such genetic variants is to quantify the absolute risk of disease conferred by the variant for an individual. This will allow valid judgements to be made about the risks and benefits of clinical interventions for the patient and family. A key assumption in our analysis is that VUS results are not reported back to patients.

Key issues or areas of uncertainty highlighted by this approach include:

Patient / genetic factors

- Limitations to current knowledge of the prevalence, spectrum and penetrance of pathogenic mutations in inherited breast cancer genes, especially those other than *BRCA1* and *BRCA2*
- The difficulty of applying estimates derived from studies in multi-case breast cancer families to the lower-risk target population undergoing WGS for a different indication
- Uncertainty about the level of absolute risk at which secondary findings in inherited breast cancer variants should be fed back, and the impact of this on an individual's clinical management
- The best approach to obtaining informed consent for the return of secondary findings of inherited breast cancer variants, given the points above

Care pathway factors

- Sub-optimal sensitivity of WGS for detecting small indels and large copy number variants or deletions
- Inconsistency between laboratories in assuring quality of WGS data and interpretation of variants
- The scale and nature of the extra work volume generated through detecting pathogenic mutations in pre-symptomatic individuals both for laboratories and clinical services
- Lack of assurance that clinical services such as clinical genetics, radiology and surgery will be able to manage this extra work volume
- Uncertainty about the cost-effectiveness of enhanced surveillance, chemoprophylaxis and / or prophylactic surgery in women from this target population who are identified with pathogenic variants in inherited breast cancer genes

At present these many uncertainties will impact on the clinical utility of actively seeking secondary findings using WGS, even for relatively well-characterised genes such as *BRCA1* and *BRCA2*. These can be addressed by appropriately planned and resourced evaluation of developing practice and research, including detailed consideration of the health economic implications. Applying the process demonstrated in this paper to other gene-disease combinations, though, is likely to highlight further gaps in knowledge and understanding around clinical validity and utility, which should be considered before offering routine testing of genes for secondary purposes in unselected populations.

4. References

1. All About The Human Genome Project [Internet]. National Human Genome Research Institute; 2015. Available from: www.genome.gov/10001772

2. Luheshi L, Raza S. Clinical whole genome analysis: delivering the right diagnosis. PHG Foundation, 2014. Available from: www.phgfoundation.org/briefing_notes/97/?

3. Pintxen W, Carmen H. Ethical issues raised by whole genome sequencing. Best Practice & Research Clinical Gastroenterology 2014; 28: 269-279.

4. Bodian DL, McCutcheon JN, Kothiyal P *et al*. Germline variation in cancer-susceptibility genes in a healthy ancestrally diverse cohort: implications for individual genome sequencing. PloS ONE 2014; 9(4):e94554.

5. Illumina TruGenome Clinical Sequencing Services [Internet]. Ilumina Clinical Services Laboratory; 2015. Available from: www.illumina.com/clinical/illumina_clinical_laboratory/trugenome-clinical-sequencing-services.html

6. Frank M, Prenzler A, Eils R, *et al*. Genome sequencing: a systematic review of health economic evidence. Health Econ. Rev. 2013; 3:1–8.

7. Green RC, Berg JS, Grody WW *et al*. ACMG recommendations for reporting of incidental findings in clinical exome and genome sequencing. Genet. Med. 2013; 15: 565-574.

8. Allyse M, Michie M. Not-so-incidental findings: the ACMG recommendations on the reporting of incidental findings in clinical whole genome and whole exome sequencing. Trends in Biotechnology 2013; 31(8): 439-441.

9. Rosenblatt DS. Who's on first in exome and whole genome sequencing? Is it the patient or the incidental findings? Molecular Genetics & Metabolism 2013; 110:1-2.

10. Burke W, Matheny Antommaria AH, Bennett R *et al*. Recommendations for returning genomic incidental findings? We need to talk! Genet. Med. 2013; 15(11): 854-859.

11. Friedman Ross L, Rothstein MA, Clayton-Wright E. Mandatory extended searches in all genome sequencing "incidental findings", patient autonomy and shared decision making. JAMA 2013; 310(4): 367-368.

12. Burton H, Zimmern R. ACMG recommendations on incidental findings: a commentary [Internet]. PHG Foundation blog; 2013. Available from: www.phgfoundation.org/blog/13713/

13. Wolf SM, Annas GJ, Elias S. Point-counterpoint. Patient autonomy and incidental findings in clinical genomics. Science 2013; 340 (6136): 1049–1050.

14. ACMG Updates recommendation on "opt out" for genome sequencing return of results [Internet]. American College of Medical Genetics and Genomics; 2014. Available from: www.acmg.net/docs/Release_ACMGUpdatesRecommendations_final.pdf

15. Anticipate and communicate: ethical management of incidental and secondary findings in the clinical, research and direct-to-consumer contexts [Internet]. Presidential Commission for the Study of Bioethical Issues; 2013. Available from: http://bioethics.gov/sites/default/files/FINALAnticipateCommunicate_PCSBI_0.pdf

16. Van El CG, Cornel MC, Borry P *et al*. Whole genome sequencing in health care. Recommendations of the European Society of Human Genetics. European Journal of Human Genetics 2013; 21: S1-S5.

17. The 100,000 Genomes Project. Genomics England Ltd. Available from: www.genomicsengland.co.uk/the-100000-genomes-project/

18. NICE Clinical Guideline 164. Familial Breast Cancer; 2013.

19. Moyer V on behalf of the US Preventive Services Task Force. Risk assessment, genetic counselling and genetic testing for *BRCA*-related cancer in women: US Preventive Services Task Force recommendation statement. Ann. Int. Med. 2014; 160: 271-281.

20. King MC, Levy-Lahad E, Lahad A. Population-based screening for *BRCA1* and *BRCA2*. 2014 Lasker Award. JAMA. Published online 8 September 2014.

21. NHS Breast Screening (BSP) Programme [Internet]. Public Health England. Available from: www.gov.uk/topic/population-screening-programmes/breast

22. Pashayan N, Duffy SW, Choudhury S *et al*. Polygenic susceptibility to prostate and breast cancer: implications for personalised screening. British Journal of Cancer 2011; 104: 1656-1663.

23. NHS Population Screening [Internet]. UK National Screening Committee. Available from: www.screening.nhs.uk/criteria

24. Burke W, Zimmern R. Moving beyond ACCE: an expanded framework for genetic test evaluation. PHG Foundation, 2007. Available from: www.phgfoundation.org/file/16270/

25. Easton D, Pharoah P, Antoniou A *et al*. Gene panel sequencing for breast cancer susceptibility – a note of caution. Manuscript in preparation. October 2014.

26. Breast cancer key facts [Internet]. Cancer Research UK. Available from: www.cancerresearchuk.org/cancer-info/cancerstats/keyfacts/breast-cancer/

27. Hollestelle A, Wasielewski M, Martens JWM *et al*. Discovering moderate-risk breast cancer susceptibility genes. Current Opinion in Genetics & Development 2010; 20:268-276.

28. Petrucelli N, Daly MB, Feldman GL *et al*. *BRCA1* and *BRCA2* hereditary breast and ovarian cancer [Internet]. Gene reviews; 1998. Available from: www.ncbi.nlm.nih.gov/books/NBK1247/#*BRCA1*.Molecular_Genetics

29. Welcsh PL, King MC. *BRCA1* and *BRCA2* and the genetics of breast and ovarian cancer. Human Molecular Genetics 2001; 10(7): 705-713.

30. Kobayashi H, Ohno S, Sasaki Y *et al*. Hereditary breast and ovarian cancer susceptibility genes (review). Oncology Reports 2013; 30: 1019-1029.

31. Figure 2: Estimated mutation detection frequencies for *BRCA1* and *BRCA2* [Internet]. *BRCA1* and *BRCA2* hereditary breast and ovarian cancer, Gene Reviews 1998. Available from: www.ncbi.nlm.nih.gov/books/NBK1247/figure/*BRCA1*.F2/?report=objectonly

32. Gomez Garcia EB, Oosterwijk JC, Timmermans M *et al*. A method to assess the clinical significance of unclassified variants in the *BRCA1* and *BRCA2* genes based on cancer family history. Breast Cancer Research 2009; 11(1).

33. Eggington JM, Burbidge LA, Roa BB *et al*. Current variant of uncertain significance rates in *BRCA1*/2 and Lynch syndrome testing (*MLH1, MSH2, MSH6, PMS2, EPCAM*). Poster. Charlotte, NC: American College of Medical Genetics Annual Clinical Genetics Meeting, 2012.

34. Fackenthal JD, Olopade OI. Breast cancer risk associated with *BRCA1* and *BRCA2* in diverse populations. Nature Reviews Cancer, 2007; 7: 937-948.

35. 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; 57: 1457-1462.

36. Nelson H, Hoyt Huffman L, Fu R *et al*. Genetic risk assessment and *BRCA* mutation testing for breast and ovarian cancer susceptibility: systematic evidence review for the US Preventive Task Force. Ann. Int. Med. 2005; 143: 362-379.

37. Christinat C, Pagani O. Practical aspects of genetic counselling in breast cancer: Lights and shadows. The Breast 2013;22: 375-382.

38. Antoniou A, Cunningham AP, Peto J *et al*. The BOADICEA model of genetic susceptibility to breast and ovarian cancers: updates and extensions. British Journal of Cancer 2008; 1-10.

39. Michailidou K, Hall P, Gonzalez-Neira A *et al*. Large-scale genotyping identifies 41 new loci associated with breast cancer risk. Nat. Genet. 2013; 45: 353-361.

40. Antoniou A, Pharoah PDP, Narod S *et al*. Average risks of breast and ovarian cancer associated with *BRCA1* and *BRCA2* mutations detected in case seris unselected for family history: a combined analysis of 22 studies. Am. J. Hum. Genet. 2003; 72: 1117-1130.

41. Chen S & Parmigiani G. Meta-analysis of *BRCA1* and *BRCA2* penetrance. J. Clin. Oncol. 2007;25(11): 1329-1333.

42. Mavaddat N, Peock S, Frost D *et al*. Cancer risks for *BRCA1* and *BRCA2* mutation carriers: results from prospective analysis of EMBRACE. J. Natl. Cancer. Inst. 2013; 105: 812-822.

43. Ford D, Easton DF, Stratton M *et al*. Genetic heterogeneity and penetrance analysis of the *BRCA1* and *BRCA2* genes in breast cancer families. Am. J. Hum. Genet. 1998; 62: 676-689.

44. Van Der Kolk D, De Bock GH, Leegte BK *et al*. Penetrance of breast cancer, ovarian cancer and contralateral breast cancer in *BRCA1* and *BRCA2* families: high cancer incidence at older age. Breast Cancer Res. Treat. 2010; 124: 643-651.

45. Xia B, Sheng Q, Nakanishi K *et al*. Control of *BRCA2* cellular and clinical functions by a nuclear partner PALB2. Molecular Cell 2006; 22: 719-729.

46. Antoniou AC, Casadei T, Heikkinen D *et al*. Breast-cancer risk in families with mutations in *PALB2*. NEJM 2014; 371(6): 497-506.

47. Janatova M, Kleibl Z, Stribma J, *et al*. The *PALB2* gene is a strong candidate for clinical testing in *BRCA1* and *BRCA2*-negative hereditary breast cancer. Cancer Epidemiol. Biomarkers Prev. 2013; 22: 2323-2332.

48. Dansonka-Mieszkowska A, Kluska A, Moes J *et al*. A novel germline *PALB2* deletion in Polish breast and ovarian cancer patients. BMC Medical Genetics 2010; 11(20):1-9.

49. Erkko H, Xia B, Nikkila J *et al*. A recurrent mutation in *PALB2* in Finnish cancer families. Nature 2007; 446: 316-319.

50. Heikkinen T, Karkkainen H, Aaltonen K *et al*. The breast cancer susceptibility mutation P*ALB2* 1592delT is associated with an aggressive tumour phenotype. Clin. Cancer Res. 2009; 15(9): 3214-22.

51. Rahman N, Seal S, Thompson D *et al*. *PALB2*, which encodes a *BRCA2*-interacting protein, is a breast cancer susceptibility gene. Nature Genetics 2007; 39(2): 165-167.

52. Foulkes WD, Ghadirian P, Akbari MR *et al*. Identification of a novel truncating *PALB2* mutation and analysis of its contribution to early-onset breast cancer in French Canadian women. Breast Cancer Res. 2007; 9(6):R83.

53. Byrnes GB, Southey MC, Hopper JL. Are the so-called low penetrance breast cancer genes *ATM, BRIP1, PALB2* and *CHEK2*, high risk for some women with strong family histories? Breast Cancer Res. 2008; 10(3): 208.

54. Antoniou AC, Casadei S, Heikkinen T *et al*. Breast-cancer risk in families with mutations in *PALB2*. NEJM 2014; 371(6): 497-506.

55. Tischkowitz M, Capanu M, Sabbaghian N *et al*. Rare germline mutations in *PALB2* and breast cancer risk: a population-based study. Hum. Mut. 2012; 33: 674-680.

56. Erkko H, Dowty J, Nikkila J *et al*. Penetrance analysis of the *PALB2* c.1592delT founder mutation. Clin. Cancer Res. 2008; 14(14):4667-4671.

57. Southey MC, Teo ZL, Dowty JG *et al*. A *PALB2* mutation associated with a high risk of breast cancer. Breast Cancer Res. 2010; 12(6):R109.

58. Gene Reviews Tumor protein *TP53* [Internet]. Gene reviews. Updated on 12 Apr 2015. Available from: www.ncbi.nlm.nih.gov/gene/7157/

59. Kast K, Krause M, Schuler M *et al*. Late onset Li Fraumeni syndrome with bilateral breast cancer and other malignancies: case report and review of the literature. BMC Cancer 2012; 12: 217.

60. Gonzalez KD, Noltner KA, Buzin CH *et al*. Beyond Li Fraumeni syndrome: clinical characteristics of families with germline *TP53* mutations. J. Clin. Oncol. 2009; 27(8): 1250-1256.

61. Lalloo F, Varley J, Ellis D *et al*. Prediction of pathogenic mutations in patients with early-onset breast cancer by family history. Lancet 2003; 361: 1101-2.

62. Custodio G, Parise GA, Kiesel Filho N *et al*. Impact of neonatal screening and surveillance for the *TP53* R337H mutation on early detection of childhood adrenocortical tumours. J. Clin. Oncol. 2013; 31(20): 2619-2627.

63. Maschiari S, Dillon DA, Rath M *et al*. Breast cancer phenotype in women with *TP53* germline mutations: a Li Fraumeni syndrome consortium effort. Breast Cancer Res. Treat. 2012; 133: 1125-30.

64. Chompret A, Brugieres L, Ronsin M *et al*. *TP53* germline mutations in childhood cancers and cancer risk for carrier individuals. British Journal of Cancer 2000; 82(12): 1932-1937.

65. Hwang SJ, Lozano G, Amos CI *et al*. Germline *TP53* mutations in a cohort with childhood sarcoma: sex differences in cancer risk. Am. J. Hum. Genet. 2003; 72(4): 975-983.

66. McCarthy M. 100,000 Genomes Project technical issues. Appendix 2. Strategic priorities in rare diseases. February 2013. Available from: www.genomicsengland.co.uk/wp-content/uploads/2013/06/GenomicsEngland_ScienceWorkingGroup_App2rarediseases.pdf

67. Meynert A, Ansari M, FitzPatrick DR *et al*. Variant detection sensitivity and biases in whole genome and exome sequencing. BMC Bioinformatics 2014; 15: 247.

68. Richards CS, Bale S, Bellissimo DB *et al*. ACMG recommendations for standards of interpretation and reporting of sequence variations: revisions 2007. Genet. Med. 2008; 10(4): 294-300.

69. Wallis Y, Payne S, McAnulty C *et al*. Practice guidelines for the evaluation of pathogenicity and the reporting of sequence variants in clinical molecular genetics. Association for Clinical Genetic Science Best Practice Guidelines June 2013.

70. Plon S, Eccles DM, Easton D *et al*. Sequence variant classification and reporting: recommendations for improving the interpretation of cancer susceptibility genetic test results. Hum. Mutat. 2008; 29(11): 1282-1291.

71. Feero WG. Clinical application of whole-genome sequencing. Proceed with care. JAMA 2014; 311(10): 1017-1019.

72. Leiden Cancer variant statistics [Internet]. Leiden Open Variation Database [Cited Oct 2014]. http://chromium.lovd.nl/LOVD2/home.php

73. Song H, Cicek MS, Dicks E *et al*. The contribution of deleterious germline mutations in *BRCA1*, *BRCA2* and the mismatch repair genes to ovarian cancer in the population. Human Molecular Genetics. 2014; 23(17): 4703-4709.

74. Ellard S, Lindsay H, Camm N *et al*. Practice guidelines for targeted next generation sequencing analysis and interpretation. ACGS Best Practice Guidelines, May 2014.

75. Rehm HL, Bale SJ, Bayrak-Toydemir P *et al*. ACMG clinical laboratory standards for next-generation sequencing. Genetics in Medicine 2013; 15(9): 733-755.

76. Illumina TruGenome Clinical Sequencing Services FAQs [Internet]. Ilumina Clinical Services Laboratory; 2015 [Cited Jan 2015]. Available from: www.illumina.com/content/dam/illumina-marketing/documents/clinical/trugenome-faq-clinical-sequencing-services-0680-2014-001.pdf

77. Warren-Gash C, Kroese M, Burton H *et al.* Implications of using whole genome sequencing to test unselected populations for high risk breast cancer genes: a modelling study. Hered. Cancer Clin. Pract. 2016; Jun 1;14:12.

78. Hall A, Finnegan T, Alberg C. Realising Genomics in Clinical Practice. PHG Foundation 2014. ISBN 978-1-907198-15-1. Available from: www.phgfoundation.org/phgpreview/cmsdemo/file/16206/

79. Dewey FE, Grove ME, Pan C *et al*. Clinical interpretation and implications of whole-genome sequencing. JAMA 2014; 311(10): 1035-1044.

80. Nelson H, Pappas M, Zakher B *et al*. Risk Assessment, Genetic Counseling, and Genetic Testing for BRCA-Related Cancer in Women: A Systematic Review to Update the U.S. Preventive Services Task Force Recommendation. Ann. Intern. Med. 2014; 160: 255-266.

5. Appendix

Appendix 1

List of genes that should be routinely examined in clinical exome or genome sequencing regardless of indication according to ACMG recommendations

Phenotype	Gene	MIM gene
Hereditary breast and ovarian cancer	*BRCA1* *BRCA2*	113705 600185
Li-Fraumeni syndrome	*TP53*	191170
Peutz-Jeghers syndrome	*STK11*	STK11
Lynch syndrome	*MLH1* *MSH2* *MSH6* *PMS2*	120436 609309 600678 600259
Familial adenomatous polyposis	*APC*	611731
MYH-associated polyposis; Adenomas, multiple colorectal; Colorectal adenomatous polyposis, autosomal recessive, with pilomatricomas	*MUTYH*	604933
Von Hippel Lindau sydrome	*VHL*	608537
Multiple endocrine neoplasia type 1	*MEN1*	613733
Multiple endocrine neoplasia type 2	*RET*	164761
Familial medullary thyroid cancer	*RET* *NTRK1*	164761 191315
PTEN hamartoma tumour syndrome	*PTEN*	601728
Retinoblastoma	*RB1*	614041
Hereditary paraganglioma-phaeochromocytoma syndrome	*SDHD* *SDHAF2* *SDHC* *SDHB*	602690 613019 602413 185470
Tuberous sclerosis complex	*TSC1* *TSC2*	605284 191092
WT1-related Wilms tumour	*WT1*	607102

Phenotype	Gene	MIM gene
Neurofibromatosis type 2	*NF2*	607379
EDS-vascular type	*COL3A1*	120180
Marfan syndrome; Loeys-Dietz syndromes; familial thoracic aortic aneurysms and dissections	*FBN1* *TGFBR1* *TGFBR2* *SMAD3* *ACTA2* *MYLK* *MYH11*	134797 190181 190182 603109 102620 600922 160745
Hypertrophic cardiomyopathy; dilated cardiomyopathy	*MYBPC3* *MYH7* *TNNT2* *TNNI3* *TPM1* *MYL3* *ACTC1* *PRKAG2* *GLA* *MYL2* *LMNA*	600958 160760 191045 191044 191010 160790 102540 602743 300644 160781 150330
Catecholinergic polymorphic ventricular tachycardia	*RYR2*	180902
Arrhythmogenic right ventricular cardiomyopathy	*PKP2* *DSP* *DSC2* *TMEM43* *DSG2*	602861 125647 125645 612048 125671
Romano-Ward Long QT syndromes types 1,2 and 3; Brugada syndrome	*KCNQ1* *KCNH2* *SCN5A*	607542 152427 600163
Familial hypercholesterolaemia	*LDLR* *APOB* *PCSK9*	606945 107730 607786
Malignant hyperthermia susceptibility	*RYR1* *CACNA1S*	180901 114208

Appendix 2

List of genes that will be routinely examined in whole genome sequencing conducted for the 100,000 Genomes Project

Phenotype	Gene	MIM gene
Hereditary non-polyposis colorectal cancer / Lynch syndrome	*MLH1*	120436
	MSH2	609309
	MSH6	600678
	PMS2	600259
Familial Adenomatous Polyposis	*APC*	611731
MYH-Associated Polyposis	*MUTYH*	604933
Hereditary breast and ovarian cancer	*BRCA1*	113705
	BRCA2	600185
Von Hippel Lindau syndrome	*VHL*	608537
Multiple Endocrine Neoplasia type 1	*MEN1*	613733
Multiple Endocrine Neoplasia type 2	*RET*	164761
Familial Medullary Thyoid Cancer	*RET*	164761
	NTRK1	191315
Retinoblastoma	*RB1*	614041
Familial Hypercholesterolaemia	*LDLR*	606945
	APOB	107730
	PCSK9	607786

phg foundation

making science work for health

About the PHG Foundation

The PHG Foundation is a pioneering independent think-tank with a special focus on genomics and other emerging health technologies that can provide more accurate and effective personalised medicine. Our mission is to make science work for health. Established in 1997 as the founding UK centre for public health genomics, we are now an acknowledged world leader in the effective and responsible translation and application of genomic technologies for health.

We create robust policy solutions to problems and barriers relating to implementation of science in health services, and provide knowledge, evidence and ideas to stimulate and direct well-informed discussion and debate on the potential and pitfalls of key biomedical developments, and to inform and educate stakeholders. We also provide expert research, analysis, health services planning and consultancy services for governments, health systems, and other non-profit organisations.

ISBN 978-1-907198-21-2

CAMBRIDGE UNIVERSITY
Health Partners

Knowledge-based healthcare

PHG Foundation
2 Worts Causeway
Cambridge
CB1 8RN
T +44 (0) 1223 761 900

www.phgfoundation.org

phg foundation

making science work for health