Gene-Panel Sequencing and the Prediction of Breast-Cancer Risk

Douglas F. Easton, Ph.D., Paul D.P. Pharoah, Ph.D., Antonis C. Antoniou, Ph.D., Marc Tischkowitz, M.D., Ph.D., Sean V. Tavtigian, Ph.D., Katherine L. Nathanson, M.D., Peter Devilee, Ph.D., Alfonso Meindl, Ph.D., Fergus J. Couch, Ph.D., Melissa Southey, Ph.D., David E. Goldgar, Ph.D., D. Gareth R. Evans, M.D., Georgia Chenevix-Trench, Ph.D., Nazneen Rahman, M.D., Ph.D., Mark Robson, M.D., Susan M. Domchek, M.D., and William D. Foulkes, M.B., B.S., Ph.D.

Advances in sequencing technology have made multigene testing, or “panel testing,” a practical option when looking for genetic variants that may be associated with a risk of breast cancer. In June 2013, the U.S. Supreme Court1 invalidated specific claims made by Myriad Genetics with respect to the patenting of the genomic DNA sequence of BRCA1 and BRCA2. Other companies immediately began to offer panel tests for breast cancer genes that included BRCA1 and BRCA2. The subsequent flourishing of gene-panel testing services (Table 1, and Table S1 in the Supplementary Appendix, available with the full text of this article at NEJM.org) has generated much interest both within the clinical genetics community and in the popular press. 2 These panels cover a total of more than 100 genes, and breast cancer is specifically mentioned as an indication for 21 of these genes. However, the fact that the technology is available does not necessarily mean that such tests are appropriate or desirable.

According to the framework proposed by the ACCE (established by the Centers for Disease Control and Prevention), genetic tests should be evaluated on the basis of the four criteria from which the name ACCE is derived: analytical validity, clinical validity, clinical utility, and ethical, legal, and social issues.3 Analytical validity refers to the degree of accuracy with which a test detects the presence or absence of a mutation. Here, however, we focus on the key question of clinical validity: Are the variants the test is intended to identify associated with disease risk, and are these risks well quantified? The validity of the risk estimates is a key determinant of the clinical utility of panel testing, which in turn should inform decisions regarding the adoption of the testing in clinical practice. We do not consider in detail who should undergo testing, what level of risk is associated with any given variant that might be considered clinically useful, or how that risk might be managed. However, broadly similar guidelines for managing the care of women with a family history of breast cancer exist in several countries (Table 2). These guidelines are based on the stratification of patients according to levels of risk and provide guidance on the identification of women to whom screening (by means of mammography or magnetic resonance imaging), risk-reducing medication, and risk-reducing surgery should be offered. These recommendations could be modified to reflect the identification of risk variants through the use of gene-panel testing. Whatever the recommendations for the management of care, the underpinnings of the guidelines should be based on reliable estimates of the risk of cancer.

Before these guidelines are developed, the appropriateness of the tests themselves needs to be considered. The determination of analytical validity for laboratory-developed diagnostic tests falls under the remit of the Clinical Laboratory Improvement Amendments (CLIA) of 1988, but neither clinical validity nor clinical utility is part of the assessment process. Therefore, whereas new drugs without clinical utility will not be approved by the Food and Drug Administration (FDA), gene-panel tests can be adopted without
any review of data regarding their clinical utility. Recent commentaries have suggested ways in which the FDA might become involved in the approval of genomic tests. Although we acknowledge the enormity of the task, we propose that a genomic test should not be offered until its clinical validity has been established. We consider below some of the key issues that need to be addressed. Others have argued that establishing clinical validity is a postmarketing pursuit, but we believe that failing to require the clinical validation of genomic tests before they are submitted for regulatory approval is likely to lead to substantial misuse of the technology.

### Table 1. Examples of Multigene Testing Panels for Breast Cancer.

<table>
<thead>
<tr>
<th>Company</th>
<th>Test</th>
<th>Website</th>
<th>Genes Included</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ambry Genetics</td>
<td>BreastNext</td>
<td><a href="http://www.ambrygen.com/tests/breastnext">www.ambrygen.com/tests/breastnext</a></td>
<td>ATM, BARD1, BRCA1, BRCA2, BRIPI, CDH1, CHEK2, MRE11A, MUTYH, NBN, NF1, PALB2, PTEN, RAD50, RAD51C, RAD51D, TP53</td>
</tr>
<tr>
<td>BreastHealth UK</td>
<td>BreastGene</td>
<td><a href="http://www.breasthealthuk.com/screening-services/genetic-testing/breastgene">www.breasthealthuk.com/screening-services/genetic-testing/breastgene</a></td>
<td>ATM, BRCA1, BRCA2, BRIPI, CDH1, CHEK2, PALB2, PTEN, STK11, TP53</td>
</tr>
<tr>
<td>Centogene</td>
<td>Breast Ovarian Cancer Panel</td>
<td><a href="http://www.centogene.com/centogene/centogene-test-catalogue.php">www.centogene.com/centogene/centogene-test-catalogue.php</a></td>
<td>ATM, BARD1, BRIPI, CDH1, CHEK2, MEN1, MLH1, MRE11A, MSH2, MSH6, MUTYH, NBN, PALB2, PMS1, PMS2, RAD50, RAD51C, RAD51D, XRCC2</td>
</tr>
<tr>
<td>Emory Genetics Laboratory</td>
<td>High Risk Breast Cancer Panel</td>
<td><a href="http://geneticslab.emory.edu/tests/MM201">http://geneticslab.emory.edu/tests/MM201</a></td>
<td>PTEN, STK11, TP53</td>
</tr>
<tr>
<td>Fulgent Diagnostics</td>
<td>Breast Ovarian Cancer NGS Panel</td>
<td><a href="http://fulgentdiagnostics.com/test/breast-ovarian-cancer-ngs-panel/">http://fulgentdiagnostics.com/test/breast-ovarian-cancer-ngs-panel/</a></td>
<td>APC, ATM, ATR, AXIN2, BAPI, BARD1, BLM, BMPR1A, BRCA1, BRCA2, BRIPI, CDH1, CDK4, CDXN2A, CHEK2, CTNNB1, EPAC, FANCC, HXXB13, MLH1, MRE11A, MSH2, MSH6, MUTYH, NBN, PALB2, PALLD, PMS2, PTEN, RAD50, RAD51, RAD51C, RAD51D, SMAD4, STK11, TP53, VHL, XRCC2, XRCC3</td>
</tr>
<tr>
<td>GeneDx</td>
<td>OncoGeneDx</td>
<td><a href="http://www.genedx.com/test-catalog/available-tests/breastovarian-cancer-panel">www.genedx.com/test-catalog/available-tests/breastovarian-cancer-panel</a></td>
<td>ATM, BARD1, BRCA1, BRCA2, BRIPI, CDH1, CHEK2, EPAC, FANCC, MLH1, MSH2, MSH6, NBN, PALB2, PMS2, PTEN, RAD51C, RAD51D, STK11, TP53, XRCC2</td>
</tr>
<tr>
<td>Illumina</td>
<td>TruSight Cancer</td>
<td><a href="http://www.illumina.com/clinical/translational_genomics/panels/kits.html">www.illumina.com/clinical/translational_genomics/panels/kits.html</a></td>
<td>94 Genes plus 287 SNPs reported to be associated with risk of breast cancer</td>
</tr>
<tr>
<td>Myriad Genetics†</td>
<td>myRisk</td>
<td><a href="http://www.myriad.com/products-services/hereditary-cancers/myrisk-hereditary-cancer/">www.myriad.com/products-services/hereditary-cancers/myrisk-hereditary-cancer/</a></td>
<td>ATM, BARD1, BRCA2, BRIPI, CDH1, CHEK2, NBN, PALB2, PTEN, RAD51C, STK11, TP53</td>
</tr>
<tr>
<td>University of Washington†</td>
<td>BROCA – Cancer Risk Panel</td>
<td><a href="http://web.labmed.washington.edu/tests/genetics/BROCA">http://web.labmed.washington.edu/tests/genetics/BROCA</a></td>
<td>AKT1, ATM, BARD1, BRCA1, BRCA2, BRIPI, CDH1, CHEK2, EPAC, FAM175A, GEN1, MRE11A, MUTYH, NBN, PALB2, PIK3CA, PTEN, RAD50, RAD51C, RAD51D, STK11, TP53, XRCC2</td>
</tr>
</tbody>
</table>

*SNP denotes single-nucleotide polymorphism.
†For Myriad Genetics and the University of Washington, only genes for which breast-cancer risk is given as an indication are listed. For a complete list, see Table S1 in the Supplementary Appendix. In several cases, the panels include additional genes, and several companies also offer larger panels. Thus, even if the primary purpose of the test is prediction of the risk of breast cancer, results will often be available (and need to be interpreted) for a larger set of genes than those listed here.
Key Issues and General Principles

Several key questions must be addressed in order to establish clinical validity. First, are variants in the gene associated with breast-cancer risk? Second, which variants, or classes of variants, are associated with risk? Third, what is the magnitude of those risks? Fourth, what methods have been used to estimate those risks? We will concentrate on the genes in which rare variants have been proposed to confer a moderate or high risk of breast cancer. For the purpose of this review, we define moderate risk as a risk of breast cancer, defined in terms of disease incidence, that is two to four times as high as that in the general population and high risk as an incidence that is more than four times as high. We leave aside the separate question of risk prediction in which profiling based on the genotyping of common polymorphisms is used (see box). We will restrict our attention to the prediction of risk in women unaffected by breast cancer, although somewhat analogous issues apply to testing in affected women. We focus on the question of breast-cancer risk, but similar considerations apply to other cancers. Indeed, some of the genes considered here also confer a predisposition to ovarian cancer, pancreatic cancer, and other cancers, and some of the available panels also include genes putatively involved in a wider range of cancers (Table 3, and Table S2 in the Supplementary Appendix). We leave aside the use of panel testing for the identification of cancer syndromes and for the management of disease in women who have cancer.

Types of Genetic Variants

Most panel testing involves identifying the coding sequences and splice junctions of the genes of interest, often in combination with alternative methods used for the detection of large genomic rearrangements. Most of the variants identified are single-base substitutions and small insertions or deletions (indels). We refer to all nonsense substitutions, frameshift indels, and variants affecting splicing as protein-truncating variants. For the large majority of genes, most of the evidence on breast-cancer risk relates to protein-truncating variants assumed to result in loss of function.

Statistical Significance and Burden Tests

It is important to establish stringent levels of statistical significance. Although it would be ideal to have specific evidence for every variant detected, most variants for which there is a suspicion of association with a high risk of disease are rare, and the sample sizes required to establish allele-specific associations with risk are so large as to make the task infeasible. Consequently, some form of burden testing is frequently used in which the association between carrying any variant in a specific class and the risk of disease is evaluated. A potential problem with this method is that it does not indicate whether any specific variant identified is associated with disease. It is often assumed that all protein-truncating variants are equally pathogenic; however, all such variants do not confer the same risks. For missense variants, the situation is even more problematic.

Strength of Statistical Evidence for Association

The issue of what constitutes appropriate levels of significance for targeted sequencing has not been extensively discussed. An exome-wide sig-
Table 2. Screening and Treatment Guidelines for Carriers of BRCA1 or BRCA2 Mutations.

<table>
<thead>
<tr>
<th>Type of Screening or Therapy</th>
<th>NCCN (United States)</th>
<th>NICE (United Kingdom)†</th>
<th>GC-HBOC–eviQ (Germany)</th>
<th>eviQ Cancer Treatments Online</th>
<th>IKNL–KiMS (Australia)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mammography</td>
<td>Recommended annually at 30–75 yr of age, or younger if woman in family has received breast-cancer diagnosis before 25 yr of age and MRI is not available</td>
<td>Recommended for consideration annually at 30–39 yr of age; recommended once yearly at 40–69 yr of age and every 3 yr at ≥70 yr of age</td>
<td>Recommended every 1–2 yr at 49–69 yr of age if breast density classified as ACR 1 or 2, with ultrasonography twice yearly‡</td>
<td>For BRCA1 carriers, recommended annually at 30–50 yr of age, with or without ultrasonography</td>
<td>Recommended annually; because risk of radiation-induced tumors is greater in young women, first mammogram recommended at 30 yr of age</td>
</tr>
<tr>
<td>MRI</td>
<td>Recommended annually at 25–75 yr of age, but earlier if younger age of onset in any family member</td>
<td>Recommended annually at 30–49 yr of age unless breast density is high, in which case should be continued until 70 yr of age</td>
<td>Recommended annually at 25–69 yr of age if breast density is classified as ACR &gt;1</td>
<td>For BRCA1 carriers, recommended annually at 30–50 yr of age, with or without ultrasonography</td>
<td>Recommended annually, starting at 25 yr of age</td>
</tr>
<tr>
<td>Preventive mastectomy</td>
<td>No definitive guideline, but &quot;degree of protection and risks&quot; should be discussed</td>
<td>No definitive guideline, but discussions of potential benefits of surgery should take current age into account</td>
<td>No definitive guideline, but &quot;degree of protection and risks&quot; should be discussed</td>
<td>If performed, recommended at ≥40 yr of age</td>
<td>Recommended at ≥25 yr of age; &lt;5% of patients are at risk of residual breast cancer</td>
</tr>
<tr>
<td>Preventive oophorectomy</td>
<td>If performed, recommended between 35 and 40 yr of age</td>
<td>No guideline</td>
<td>Salpingo-oophorectomy recommended at approximately 40 yr of age for BRCA1 carriers and 45 yr of age for BRCA2 carriers</td>
<td>If performed, recommended at ≥40 yr of age</td>
<td>If performed, recommended at ≥35 yr of age for BRCA1 carriers and ≥40 yr of age for BRCA2 carriers</td>
</tr>
<tr>
<td>Oral contraceptive</td>
<td>No clear directive</td>
<td>No clear directive</td>
<td>No clear directive</td>
<td>Combination oral contraceptive not contraindicated</td>
<td>No clear directive; recommended that nonsystemic form of contraception could be discussed</td>
</tr>
<tr>
<td>Chemoprevention</td>
<td>No clear directive</td>
<td>Provision of tamoxifen recommended for women at high risk of breast cancer; but BRCA1 vs. BRCA2 status not discussed</td>
<td>No guideline</td>
<td>Recommendation to consider with professional on individualized basis</td>
<td>No guideline</td>
</tr>
</tbody>
</table>
Significance level of $P<2.5\times 10^{-6}$ is often used for whole-exome studies (calculated on the basis of a Bonferroni correction for approximately 20,000 genes). Since most genes associated with susceptibility to breast cancer are involved in DNA repair (a class involving fewer than 500 genes), more liberal significance levels (on the order of $P<0.0001$) might be appropriate for genes in this pathway. The use of Bayesian arguments leads to similar thresholds (see the Methods section in the Supplementary Appendix). Although these significance thresholds may be appropriate for a single burden test, more stringent thresholds would be required for calculations involving individual variants. A related question is the precision of the risk estimate. It is clearly undesirable to give a patient an estimate of risk that may be subject to substantial change when additional data are acquired. For the purposes of this review, we consider it to be likely that a given risk will be above (or below) a certain threshold if the 90% confidence limit on the risk estimate exceeds (or is less than) the threshold.

**Definition of Risk**

Our estimates are presented primarily in terms of average relative risks. We recognize that for purposes of counseling, absolute estimates of risk (projected over a few years or a lifetime) are more useful. However, most studies report estimates of relative risk rather than absolute risk, and absolute risks are more strongly influenced by risk factors for breast cancer, such as a family history of breast cancer, age at menopause, and breast density on mammography. In the case of a rare variant conferring a relative risk of 2 or 4, the corresponding absolute risks of breast cancer would be approximately 18% and 32%, respectively, by the time a patient reached 80 years of age (according to recent U.K. incidence rates), in the absence of other causes of death. These risks approximately correspond to the definitions of moderate and high risk familiar to the clinical genetics community.

It follows that the identification of a variant conferring a relative risk higher than 4, in the absence of any other data, can place a woman in the high-risk category. In contrast, a variant conferring a relative risk of 2 to 4 will place a woman in the high-risk category only if her risk is increased by other factors. For some genes (notably, $BRCA1$, $CHEK2$, and $ATM$), there is
Table 3. Genes for Which an Association between Protein-Truncating Variants and Breast-Cancer Risk Has Been Established.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Magnitude of Relative Risk Associated with Truncating Variants*</th>
<th>Risk Associated with Missense Variants†</th>
<th>Estimated Relative Risk (90% CI)‡</th>
<th>P Value</th>
<th>Absolute Risk by 80 Yr of Age§</th>
<th>Comments</th>
<th>Other Associated Cancers</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>BRCA1</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>11.4</td>
<td>75</td>
<td>Estimates are based on the BOADICEA model for a woman born in 1960</td>
<td>Ovary</td>
<td>Antoniou et al.,10 Lee et al.,11 Chen and Parmigiani,12 Mavaddat et al.13</td>
</tr>
<tr>
<td>BRCA2</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>11.7</td>
<td>76</td>
<td>Estimates are based on the BOADICEA model for a woman born in 1960; p.Lys3326Ter in the carboxyl terminus is associated with a lower increase in risk</td>
<td>Ovary, prostate, pancreas</td>
<td>Antoniou et al.,10 Lee et al.,11 Chen and Parmigiani,12 Mavaddat et al.13</td>
</tr>
<tr>
<td>TP53</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>105 (62–165)</td>
<td>Most published risk estimates are subject to ascertainment bias</td>
<td>Childhood sarcoma, adrenocortical carcinoma, brain tumors</td>
<td></td>
<td>Hisada et al.,14 Hwang et al.15</td>
</tr>
<tr>
<td>PTEN</td>
<td>Unknown</td>
<td>Unknown</td>
<td>Yes</td>
<td>No reliable estimate‖</td>
<td>Published risk estimates are subject to ascertainment bias</td>
<td>Thyroid, endometrial cancer</td>
<td></td>
<td>Bubien et al.,16 Tan et al.17</td>
</tr>
<tr>
<td>CDH1</td>
<td>Likely</td>
<td>Unknown</td>
<td>Unknown</td>
<td>6.6 (2.2–19.9)</td>
<td>0.004</td>
<td>Specific to lobular breast cancer</td>
<td>Diffuse gastric cancer</td>
<td>Pharoah et al.18</td>
</tr>
<tr>
<td>STK11</td>
<td>Unknown</td>
<td>Unknown</td>
<td>Unknown</td>
<td>No reliable estimate‖</td>
<td>Published risk estimates are subject to ascertainment bias</td>
<td>Colon, pancreas, ovarian sex cord–stromal tumors</td>
<td></td>
<td>Hearle et al.19</td>
</tr>
<tr>
<td>NF1</td>
<td>Likely</td>
<td>Unlikely</td>
<td>Unknown</td>
<td>2.6 (2.1–3.2)</td>
<td>2.3×10⁻¹³</td>
<td>Estimates are based on cohort studies of patients with neurofibromatosis type I ††</td>
<td>Malignant tumors of peripheral nerve sheath, brain, central nervous system</td>
<td>Madanikia et al.,20 Seminog and Goldacre21</td>
</tr>
<tr>
<td>PALB2</td>
<td>Likely</td>
<td>Unknown</td>
<td>Unknown</td>
<td>5.3 (3.0–9.4)</td>
<td>4×10⁻¹⁰</td>
<td>Estimates are based on a meta-analysis of published case–control and family studies</td>
<td>Pancreas</td>
<td>Antoniou et al.,22 Heikkinen et al.,23 Rahman et al.,24 Erkko et al.25</td>
</tr>
</tbody>
</table>
The p.Val2424Gly variant is associated with higher risk than truncating variants.

**CHEK2**
- Likely: Yes
- Unlikely: Yes
- **2.8 (2.2–3.7)**
- **5×10⁻¹¹**
- Pancreas

Pancreas Renwick et al.,

**CHEK2**

**CHEK2**
- Likely: Yes
- Unlikely: Yes
- **2.7 (1.9–3.7)**
- **5×10⁻⁷**
- Lung, although p.Ile157Thr is associated with reduced risk

Renwick et al.,

**CHEK2**

Most data for truncating variants are limited to the variant c.1100delC; p.Ile157Thr is associated with an increase in risk that is 1.3 times as high as in the general population

Most data for truncating variants are limited to the variant c.1100delC; p.Ile157Thr is associated with an increase in risk that is 1.3 times as high as in the general population

Janin et al.,

27 The p.Val2424Gly variant is associated with higher risk than truncating variants

Renwick et al.,

**CHEK2**

**CHEK2**
- Likely: Yes
- Unlikely: Yes
- **3.0 (2.6–3.5)**
- **8×10⁻³⁷**
- Lung, although p.Ile157Thr is associated with reduced risk

Most data for truncating variants are limited to the variant c.1100delC; p.Ile157Thr is associated with an increase in risk that is 1.3 times as high as in the general population

Renwick et al.,

**CHEK2**

**CHEK2**
- Likely: Yes
- Unlikely: Yes
- **2.8 (2.2–3.7)**
- **5×10⁻¹¹**
- Pancreas

Pancreas Renwick et al.,

**CHEK2**

Most data for truncating variants are limited to the variant c.1100delC; p.Ile157Thr is associated with an increase in risk that is 1.3 times as high as in the general population

Pancreas Renwick et al.,

**CHEK2**

Most data for truncating variants are limited to the variant c.1100delC; p.Ile157Thr is associated with an increase in risk that is 1.3 times as high as in the general population

Renwick et al.,
Table 4. Study Designs for Estimating the Risks Associated with Rare Variants.

<table>
<thead>
<tr>
<th>Method</th>
<th>Description</th>
<th>Advantages</th>
<th>Disadvantages</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population-based</td>
<td>Screening is conducted for variants in unselected cases of disease and population-matched controls</td>
<td>Provides direct estimates of the relative risk (odds ratio) and is not biased by other familial factors</td>
<td>Must be very large since variants are typically rare; biases arise if controls are not appropriately population-matched (there are large differences in allele frequency among populations); Requires that the same assay techniques are used for cases and controls to provide valid tests and estimates, typically with screening of the full coding sequences in all cases and controls; large biases may arise if only the variants identified in the cases are tested in the controls Requires that the same assay techniques are used for cases and controls to provide valid tests and estimates, typically with screening of the full coding sequences in all cases and controls; large biases may arise if only the variants identified in the cases are tested in the controls</td>
<td></td>
</tr>
<tr>
<td>case–control</td>
<td></td>
<td></td>
<td></td>
<td>CHEK²¹</td>
</tr>
<tr>
<td>Family-based</td>
<td>In these case–control studies, cases are enriched by family histories</td>
<td>Improves power because of the higher frequency of variants in familial cases</td>
<td>Is subject to biased risk estimates Efforts to correct bias depend on use of additional assumptions about the modifying effects of other familial factors</td>
<td>CHEK²⁰</td>
</tr>
<tr>
<td>case–control</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kin cohort</td>
<td>Data on cancer occurrence in relatives of carriers in population-based series are used to estimate risks with maximum-likelihood methods</td>
<td>Provides estimates without the need to screen controls; genotype data in relatives can be incorporated but are not required</td>
<td>Is limited by the accuracy of the family history; risks may be overestimated if familial factors are not accounted for</td>
<td>BRCA1/2, PALB²¹,²²</td>
</tr>
<tr>
<td>Segregation in</td>
<td>Can be applied in families that have been selected for a strong family history; controls are not required</td>
<td></td>
<td>Requires samples on multiple persons from the same family; power is typically very limited</td>
<td>CHEK²⁰</td>
</tr>
<tr>
<td>families</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prospective cohort</td>
<td>Provides direct estimates of absolute risk</td>
<td></td>
<td>Requires long-term investment; is prohibitively large except in the case of high-risk variants; risk estimates may be altered by interventions (e.g., prophylactic surgery); risk estimates are affected by other familial factors</td>
<td>BRCA1/2¹³</td>
</tr>
</tbody>
</table>
evidence that the rate ratio declines with age. The published overall relative risk estimates can thus provide a misleading estimate of lifetime risk. Ideally, age-specific estimates are required, but the data available on risks for older women are often limited.

**STUDY DESIGN**

Appropriate study design is critical for both the identification of disease-associated alleles and the derivation of reliable risk estimates. Several study designs are available (Table 4). The use of case–control studies for risk estimation involving rare variants can be problematic; family-based methods, including kin–cohort designs and cosegregation analysis, provide alternatives, but these methods also have pitfalls. Furthermore, many studies are based on a few variants that are restricted to specific populations; although it is generally assumed that the risk estimates associated with different truncating variants observed in other populations are similar, it is usually impossible to test this assumption.

**OVERESTIMATION OF RISK**

The problems of publication bias, in which negative studies are not published, and winner's curse, whereby an initial study identifying an association tends to overestimate the risk, should be noted. Furthermore, many gene-discovery studies oversample for early-onset cases of disease or cases with a family history. This approach improves power but leads to seriously biased risk estimates unless the ascertainment is allowed for in the analysis. Moreover, risk estimates based on data from highly selected families may not reflect the true “average” risk for all carriers of pathogenic variants, because such biased sampling results in a selection of individuals that are not random with respect to other modifiers of risk.

**EVIDENCE OF ASSOCIATION FOR SPECIFIC GENES**

Here we review several genes for which some evidence of an association with breast cancer has been reported. A summary of the genes for which an association with breast cancer has, in our view, been established is given in Table 3. See Table S2 in the Supplementary Appendix for a list of genes for which an association with breast cancer has been suggested but not established and Table S3 for a summary of the studies used to derive estimates of breast-cancer risk. The Methods section in the Supplementary Appendix summarizes the methods used to derive summary estimates of risk.

**BRCA1 AND BRCA2**

The clinical validity and utility of testing for variants in BRCA1 and BRCA2 are well established. There is overwhelming evidence that most protein-truncating variants in these genes are associated with a high risk of breast cancer and other cancers. Even among protein-truncating variants, however, variant-specific differences in risk have been observed. Furthermore, a polymorphic nonsense variant at the carboxyl terminus of BRCA2, p.Lys3326Ter, has been reported to be associated with a relative risk of breast cancer of 1.4 (90% confidence interval [CI], 1.2 to 1.7), which is substantially lower than the risks conferred by more proximal truncating variants (Table 3).

**TP53, CDH1, PTEN, STK11, AND NF1**

Mutations in TP53, CDH1, PTEN, STK11, and NF1 cause pleiotropic tumor syndromes in which breast cancer is only one feature. Germline mutations in TP53 (both protein-truncating and missense mutations) are responsible for the Li–Fraumeni syndrome, in which carriers are predisposed to childhood sarcomas, brain tumors, adrenocortical carcinoma, and other rare cancers, in addition to breast cancer. Although the association with breast cancer is not controversial, reliable estimates of risk are lacking; most studies are based on pedigrees in which family members have features of the Li–Fraumeni syndrome and thus are subject to ascertainment bias. However, a study based on carriers of a TP53 mutation identified through probands with childhood sarcoma has also reported a high risk of breast cancer. Similar ascertainment biases apply to mutations in PTEN and STK11. Mutations in PTEN are associated with the Cowden syndrome, in which breast cancer is a characteristic of the clinical phenotype, and mutations in STK11 are associated with the Peutz–Jeghers syndrome and an increased risk of breast cancer. Protein-truncating variants in CDH1, which are known to be associated with diffuse-type gastric cancer, are also thought to be as-
associated with an increased risk of breast cancer (specifically, the lobular subtype), with a reported relative risk of 6.6 (90% CI, 2.2 to 19.9; P = 0.004). Recent cohort studies have reported an elevated risk of breast cancer in women with neurofibromatosis type 1 (odds ratio, 2.6; 90% CI, 2.1 to 3.2).

**PALB2, CHEK2, ATM, NBN, and Related Genes**

There is strong evidence that protein-truncating variants in four other genes involved in DNA repair confer an increased risk of breast cancer. Among these genes, mutations in PALB2 appear to confer the highest risks. A large family-based study estimated the risk of breast cancer to be approximately six times as high among carriers as compared with noncarriers, although two case–control studies based on the Finnish founder variant, c.1592delT, estimated somewhat lower risks. In a meta-analysis of these estimates the combined relative risk was 5.3 (90% CI, 3.0 to 9.4). Thus, although PALB2 mutations may fall into the high-risk category (in which the risk of cancer is more than four times as high as that in the general population), the confidence limits are too wide to be certain. Most of the data for CHEK2 relate to the c.1100delC variant, which is found fairly frequently in Northern European populations. On the basis of two large case–control analyses, we calculated an estimated relative risk of breast cancer of 3.0 (90% CI, 2.6 to 3.5). Truncating variants in ATM have been evaluated in both case–control studies (with selected cases) and cohort studies of relatives of patients with ataxia–telangiectasia. In a meta-analysis of the three largest cohort studies of relatives of patients with ataxia–telangiectasia, the estimated relative risk of breast cancer was 2.8 (90% CI, 2.2 to 3.7; P = 4.7×10−13), a value similar to that for truncating variants in CHEK2.

In NBN, one protein-truncating variant, c.657del5, is sufficiently common in some Eastern European populations to allow its evaluation in a case–control study. A meta-analysis of 10 studies reported strong evidence of an association with breast-cancer risk for this variant (summary relative risk, 2.7; 90% CI, 1.9 to 3.7; P = 5×10−7). More limited evidence is available for two other DNA-repair genes, MRE11A and RAD50, which encode proteins that form an evolutionarily conserved complex with NBN.

Mutations in three other DNA repair genes, RAD51C, RAD51D, and BRIPI, have shown clear evidence of an association with ovarian cancer. However, in each case, the evidence for association with breast cancer is limited. Recent exome studies and targeted sequencing studies have suggested that breast cancer is associated with deleterious variants in FANCC, FANCM, and XRCC2. In none of these instances, however, does the evidence reach the threshold level (P < 0.0001) that we propose for DNA-repair genes. The recent findings of deleterious mutations in RECQL in women with a strong family history of breast cancer, however, suggests that this gene confers susceptibility to breast cancer.

**Other Genes**

The panels currently marketed for the prediction of risk of cancer contain many other genes, most of which have been included by virtue of their relevance to rare mendelian cancer syndromes. Variants in some of these genes may also be associated with breast cancer. Mutations in DNA mismatch-repair genes (MLH1, MSH2, MSH6, and PMS2) may be associated with breast cancer, but in a recent review, Win et al. concluded that the evidence was equivocal. It has also been suggested that MUTYH variants that confer a predisposition to polyposis colorectal cancer may confer a predisposition to breast cancer, but a recent case–control study reported no association. Another recent study suggested that carriers of MEN1 mutations may be at increased risk for breast cancer. A recent case–control study has reported an association between rare variants in PPM1D and breast cancer. However, this association does not reach our proposed significance threshold, and, in addition, the sequence variants are observed as mosaics in lymphocytes and are not inherited. There is currently no clear evidence of an association between breast cancer and any other gene.

**Missense Variants**

With the exception of TP53, the assessment of the risk of breast cancer from missense variants is much more problematic than it is for protein-truncating variants. Some missense variants in specific domains of BRCA1 and BRCA2 confer high risks of breast and ovarian cancer, but the
The majority do not. For these genes, algorithms based on conservation, pedigree data, and analysis of tumor subtype can be used to predict the pathogenicity of some variants. Similar considerations may apply to ATM and CHEK2 — missense variants falling in key functional domains and at positions that show a high degree of species conservation are more likely to be associated with increased risk. However, even for BRCA1 and BRCA2, the breast-cancer risk associated with the large majority of missense variants remains unknown; such variants are referred to as variants of unknown significance. Moreover, clearly pathogenic missense variants need not be associated with the same risk as truncating variants. For example, the CHEK2 missense variant p.Ile157Thr confers a lower risk of breast cancer than the CHEK2 c.1100delC truncating variant, whereas ATM p.Val2424Gly appears to be associated with a higher risk of breast cancer than truncating variants (8.0; 90% CI, 2.8 to 22.5; \( P = 0.0005 \)). A more systematic approach to this problem would involve defining risks on the basis of variant classes that are defined through prediction algorithms based on in silico data. However, even though existing data provide good evidence that missense variants falling at highly conserved positions in several genes confer disease risk, and that such variants may make an important contribution to the heritability of breast cancer, no system has been established for use in the classification of variants that would allow such estimates of risk to be used clinically.

RISK MODIFIERS AND ABSOLUTE RISKS

For the purposes of genetic counseling, relative risks need to be converted into absolute risks. For an “average” mutation carrier, absolute risks can be calculated in a straightforward manner by combining the estimated relative risk with population incidence rates. The results are illustrated in Figure 1 for carriers of mutations in PALB2 and CHEK2.

However, the calculation of the absolute risk associated with a given variant must also account for the risk associated with other genetic factors, lifestyle, and family history. There is strong evidence that the absolute risk of breast cancer in carriers of BRCA1, BRCA2, PALB2, and
CHEK2 mutations is higher among women with a strong family history of breast cancer.\(^{10,21,30,70}\) It has also been shown that the absolute risk of breast cancer in carriers of BRCA1 and BRCA2 mutations depends on the risks associated with their single-nucleotide–polymorphism (SNP) profile.\(^{71}\) A broader question is that of how the risks associated with genetic variants should be combined with risk factors associated with lifestyle. Several studies indicate that the risks associated with common SNPs and other risk factors combine in a multiplicative rather than an additive fashion,\(^{72-74}\) and it would be reasonable to assume that rare variants combine with other risk factors in a similar manner. The evidence regarding the combined effects of genetic and lifestyle factors is both limited and conflicting for variants in BRCA1 and BRCA2,\(^{75}\) and no evidence is available for other genes. In addition, absolute risks need to be adjusted for competing risks in analyses of mortality, a factor that may be important in to our understanding of genes associated with cancers other than breast cancer.

Almost all the available data relate to women of European ancestry. At present, it is unclear whether the available estimates of relative risk can be safely extrapolated to women of other ancestries or to populations with different incidences of breast cancer.

**Conclusions**

We have discussed some of the difficulties of assigning risk to rare variants and reviewed the genes for which the evidence of association with breast cancer is sufficiently robust to be incorporated into personalized risk prediction. Variants that are predicted to truncate BRCA1 and BRCA2 (together with a subset of missense variants) confer a high risk of breast cancer; PALB2 and perhaps PTEN may also fall in this category, but the evidence is insufficient to place them confidently in the category of high risk rather than moderate risk. For TP53, both missense and protein-truncating variants are associated with substantially increased risks of breast cancer. Genes that fall into the category of moderate risk (for which fully deleterious mutations confer a risk of breast cancer that is two to four times as high as that in the general population) include CHEK2, ATM, and NF1. There is clear evidence for an association with risk of cancer for STK11, CDH1, and NBN, but the risk estimates are too imprecise for categorization. Estimates of risk for PTEN, STK11, and CDH1 are derived entirely from studies of selected patients identified through specialized clinics and may be seriously overestimated. We found insufficient evidence to establish any other genes as conferring a predisposition for breast cancer and would caution against their use in the prediction of breast-cancer risk. As the costs of sequencing decline, it is inevitable that the use of gene-panel testing, and indeed whole-exome and whole-genome sequencing, will become widespread. Therefore, there is an urgent need for much larger, well-designed population- and family-based studies in diverse populations that will provide reliable estimates of risk for the purpose of counseling. The systematic collection of data from ongoing use of panel testing linked to the epidemiologic and clinical data may also make an important contribution. Other genes that convey susceptibility to breast cancer (and perhaps rarer variants in noncoding sequences) will probably be identified and may be added to genetic-testing panels. Panel testing can make a useful contribution to prediction of a woman’s risk of breast cancer, but end users need to be aware of the limitations of these panels.

Supported by grants from Cancer Research UK (A11174, to Dr. Antoniou), the National Institutes of Health (CA116167, CA176785, and CA192393, to Dr. Couch), the Breast Cancer Research Foundation (to Drs. Couch, Nathanson, and Robson), Susan G. Komen (to Drs. Foulkes, Chenevix-Trench, and Domchek), the Cancer Research Society–Quebec Breast Cancer Foundation (to Dr. Foulkes), the National Health and Medical Research Council (to Dr. Chenevix-Trench), and the Basser Research Center (to Dr. Domchek).

Disclosure forms provided by the authors are available with the full text of this article at NEJM.org.

We thank Paul James, Ingrid Winship, and Christi van Asperen for assistance with the summary of clinical guidelines.

From the Departments of Public Health and Primary Care (D.F.E., P.D.P.P., A.C.A.), Oncology (D.F.E., P.D.P.P.), and Medical Genetics (M.T.), University of Cambridge, Cambridge, the Centre for Genomic Medicine, Institute of Human Development, Manchester Academic Health Science Centre, University of Manchester and St. Mary’s Hospital, Manchester (D.G.R.E.), and the Division of Genetics and Epidemiology, Institute of Cancer Research, London (N.R.) — all in the United Kingdom; the Departments of Oncological Sciences (S.V.T.) and Dermatology (D.E.G.), Huntsman Cancer Institute, University of Utah School of Medicine, Salt Lake City; the Basser Research Center for BRCA and Abramson Cancer Center, Perelman School of Medicine at the University of Pennsylvania, Philadelphia.
This article was published on May 27, 2015, at NEJM.org.

3. Haddow JE, Palomaki GE. ACCE: a model process for evaluating data on emerging genetic tests. In: Khoury MJ, Little J, Parkville, VIC (M.S.), and the QIMR Berghofer Medical Research Institute, Herston, QLD (G.C.-T.) — both in Australia; the Clinical Genetics Service, Division of Solid Tumor Oncology, Department of Medicine, Memorial Sloan Kettering Cancer Center, New York (M.R.); and the Program in Cancer Genetics, Departments of Human Genetics and Oncology, the Lady Davis Institute for Medical Research, and the Research Institute of the McGill University Health Center, McGill University, Montreal (W.D.F.).

3. Haddow JE, Palomaki GE. ACCE: a model process for evaluating data on emerging genetic tests. In: Khoury MJ, Little J, Parkville, VIC (M.S.), and the QIMR Berghofer Medical Research Institute, Herston, QLD (G.C.-T.) — both in Australia; the Clinical Genetics Service, Division of Solid Tumor Oncology, Department of Medicine, Memorial Sloan Kettering Cancer Center, New York (M.R.); and the Program in Cancer Genetics, Departments of Human Genetics and Oncology, the Lady Davis Institute for Medical Research, and the Research Institute of the McGill University Health Center, McGill University, Montreal (W.D.F.).

This article was published on May 27, 2015, at NEJM.org.

3. Haddow JE, Palomaki GE. ACCE: a model process for evaluating data on emerging genetic tests. In: Khoury MJ, Little J, Parkville, VIC (M.S.), and the QIMR Berghofer Medical Research Institute, Herston, QLD (G.C.-T.) — both in Australia; the Clinical Genetics Service, Division of Solid Tumor Oncology, Department of Medicine, Memorial Sloan Kettering Cancer Center, New York (M.R.); and the Program in Cancer Genetics, Departments of Human Genetics and Oncology, the Lady Davis Institute for Medical Research, and the Research Institute of the McGill University Health Center, McGill University, Montreal (W.D.F.).

This article was published on May 27, 2015, at NEJM.org.

3. Haddow JE, Palomaki GE. ACCE: a model process for evaluating data on emerging genetic tests. In: Khoury MJ, Little J, Parkville, VIC (M.S.), and the QIMR Berghofer Medical Research Institute, Herston, QLD (G.C.-T.) — both in Australia; the Clinical Genetics Service, Division of Solid Tumor Oncology, Department of Medicine, Memorial Sloan Kettering Cancer Center, New York (M.R.); and the Program in Cancer Genetics, Departments of Human Genetics and Oncology, the Lady Davis Institute for Medical Research, and the Research Institute of the McGill University Health Center, McGill University, Montreal (W.D.F.).


DOI: 10.1056/NEJMsr1501341

Copyright © 2015 Massachusetts Medical Society.