POSITION STATEMENT

Towards controlled terminology for reporting germline cancer susceptibility variants: an ENIGMA report

Amanda B Spurdle,[1] Stephanie Greville-Heygate,[2] Antonis C Antoniou,[3] Melissa Brown,[4] Leslie Burke,[4] Miguel de la Hoya,[5] Susan Domchek,[6] Thilo Dörk,[7] Helen V Firth,[8] Alvaro N Monteiro,[9] Arjen Mensenkamp,[10] Michael T Parsons,[1] Paolo Radice,[11] Mark Robson,[12] Marc Tischkowitz,[13] Emma Tudini,[1] Clare Turnbull,[14,15] Maaike PG Vreeswijk,[16] Logan C Walker,[17] Sean Tavtigian,[18,19] Diana M Eccles[2]

ABSTRACT

The vocabulary currently used to describe genetic variants and their consequences reflects many years of studying and discovering monogenic disease with high penetrance. With the recent rapid expansion of genetic testing brought about by wide availability of high-throughput massively parallel sequencing platforms, accurate variant interpretation has become a major issue. The vocabulary used to describe single genetic variants in silico, in vitro, in vivo and as a contributor to human disease uses terms in common, but the meaning is not necessarily shared across all these contexts. In the setting of cancer genetic tests, the added dimension of using data from genetic sequencing of tumour DNA to direct treatment is an additional source of confusion to those who are not experienced in cancer genetics. The language used to describe variants identified in cancer susceptibility genetic testing typically still reflects an outdated paradigm of Mendelian inheritance with dichotomous outcomes. Cancer is a common disease with complex genetic architecture; an improved lexicon is required to better communicate among scientists, clinicians and patients, the risks and implications of genetic variants detected. This review arises from a recognition of, and discussion about, inconsistencies in vocabulary usage by members of the ENIGMA international multidisciplinary consortium focused on variant classification in breast-ovarian cancer susceptibility genes. It sets out the vocabulary commonly used in genetic variant interpretation and reporting, and suggests a framework for a common vocabulary that may facilitate understanding and clarity in clinical reporting of germline genetic tests for cancer susceptibility.

BACKGROUND

The Evidence-Based Network for the Interpretation of Germline Mutant Alleles (ENIGMA) consortium is an international effort focused on determining the clinical significance of variants in breast-ovarian cancer genes. In addition, ENIGMA provides expert opinion to global classification and database initiatives, notably ClinGen (Clinical Genome Resource; https://www.clinicalgenome.org/), ClinVar (https://www.ncbi.nlm.nih.gov/clinvar/) and the BRCA-Exchange (http://brcaexchange.org/). ENIGMA also explores optimal avenues of communication of such information at the provider and patient level. Importantly, most members (65%) conduct research and clinical activities in a language other than English (see online supplementary text).

ENIGMA research initially focused on improvement of methods to classify *BRCA1* (MIM113705) and *BRCA2* (MIM600185) variants associated with typical 'high' risk of cancer,[1] with subsequent investigations identifying *BRCA1/2* variants associated with demonstrably lower cancer risks.[2 3] The inclusion of multicancer syndrome and novel breast-ovarian cancer susceptibility genes on research and commercial cancer gene panels has expanded the scope of ENIGMA investigations. Four consecutive ENIGMA consortium meetings have included dedicated time to discuss appropriate terminology for describing genetic variants, and their relationship to risk of different cancer types, and implications for clinical management. In particular, as genetic test ordering has moved outside the traditional hereditary cancer clinic setting into mainstream oncology, concern has been raised regarding misinterpretation of variant pathogenicity descriptions—even for well-characterised genes like *BRCA1/2*.[4]

ENIGMA members spanning all ENIGMA working groups have developed a document that provides an overview of different terms used in scientific and clinical reports, and by relevant international bodies, to describe various aspects of sequence variation in cancer predisposition genes. This exercise revealed alternative usage for many terms, interchangeable use of terms, and the potential for misinterpretation of the actionability of variants. We sought feedback from the general ENIGMA membership, by circulation of a draft discussion document and presentation at three consecutive consortium meetings, regarding their views on which terms may be most appropriate for promotion as preferred terminology in ENIGMA documentation, research projects and manuscripts. Discussions highlighted in particular the complexities of describing variant association with cancer risk

in the context of multigene panel tests. Namely, that such tests may include genes for which 'pathogenic' variants are associated with varying levels of risk for different cancer types, and where, even for specific genes with well-established hereditary cancer risk profiles, some variants may be associated with altered cancer penetrance compared with the 'average pathogenic' variant for that gene. Different terms in use were considered by ENIGMA members attending the June 2018 Consortium Meeting, to reach consensus about the least ambiguous terms for clinical reporting. We provide some general recommendations for terminology to describe cancer susceptibility gene variation and its relationship to risk. We also propose a multitier structure for reporting cancer susceptibility variants, to improve the understanding of level of cancer risk associated with an identified variant and appropriate clinical actionability given patient presentation.

The need for standardised terminology and definitions for describing sequence variation, focused on inherited variants

Online supplementary table 1 summarises terms used to describe sequence variants, and their association with or relevance to disease, and to patient clinical management. The information was derived from a combination of knowledge from the literature, usage in verbal and written project reporting across ENIGMA, in clinical reports generated or viewed by ENIGMA members and documentation/terms described by the Human Variome Project (HVP; http://www.humanvariomeproject.org), ClinVar (https://www.ncbi.nlm.nih.gov/clinvar/) and International Society for Gastrointestinal Hereditary Tumours (InSiGHT; https://www.insight-group.org/). The content was presented to ENIGMA members at several consecutive consortium meetings, and also circulated in document form, to invite feedback and additions. While not claiming to be an exhaustive list of terms and their meanings, it is clear that a single term/phrase can be used to describe different aspects relating to a variant (different intent), and that multiple terms can describe just one aspect (same usage). In some instances, differences in terminology appeared to depend on the field of research, and the context in which a variant is identified. Notably, the term 'pathogenic variant' is used to describe a germline disease-causing variant in a Mendelian disease gene classified according to criteria from the American College of Medical Genetics and Genomics/Association for Molecular Pathology (ACMG/AMP)[5] or International Agency for Research on Cancer (IARC).[6] It has also been described as a 'sequence variant that contributes mechanistically to disease but is not necessarily fully penetrant, that is, may not be sufficient in isolation to cause disease' in the context of assessing support of disease causality of variants identified by high-throughput sequencing.[7] Moreover, a germline 'pathogenic variant' considered causal for disease risk is commonly termed a 'mutation' in the historical and even current literature, and in the medical management (National Comprehensive Cancer Network, www.nccn.org; National Institute for Health and Care Excellence, https://www.nice.org.uk; EviQ, https://www.eviq.org.au) and research setting.[8] However, 'mutation' can refer to any permanent change in DNA sequence (irrespective of frequency or disease-causing potential), and 'mutation' is used almost exclusively to describe somatic variation in the context of tumourigenesis. Indeed, the interface of the Clinical Interpretation of Variants in Cancer (CIViC) knowledgebase[9] describes variants for a specific gene using the term 'mutation', with additional qualifications, for example, for *TP53* (MIM191170), the qualifiers include: deleterious, DNA binding domain, truncating. To add to the complexity, the Leiden Open Variation Database (LOVD)

freeware database software,[10] promoted widely for sharing and curation of (germline) disease gene variants, describes the equivalent of variant pathogenicity as 'variant effect'. The most current version LOVD3 prescribes the terms 'affects function' instead of 'pathogenic', and the following terms for four other pathogenicity classes: 'probably affects function', 'unknown (or effect on function not known)', 'probably does not affect function (or probably no functional effect)' and 'does not affect function (or no functional effect)'.

Furthermore, feedback from ENIGMA consortium members indicated there was varied perception of the level of risk association and clinical actionability for variants described as 'benign' or 'not pathogenic', terms put forward by the ACMG/AMP[5] and IARC[6] classification schemes, respectively, to indicate that a variant is not clinically actionable for patient management. Also, the distinction between a variant described as uncertain (ACMG/AMP and IARC—reviewed and insufficient or conflicting evidence regarding pathogenicity) versus unclassified (not yet assessed)[11] was poorly recognised.

In addition, we separately documented terms used to describe output from some more commonly used bioinformatic prediction tools (table 1), since results from bioinformatic analysis are almost always included in clinical test reports. Such bioinformatic predictions are generally defined without reference to clinical information, are often binary and are intended to be included as one of *several* points of information used to arrive at a final variant classification. Nevertheless, we identified several possibilities for misinterpretation of bioinformatic output terms as a 'final' variant classification. The PolyPhen2 tool[12] uses the term 'benign' to describe variants with no/little predicted effect on protein function—the same as the ACMG/AMP term for a variant that is not considered important for diagnosis/risk/patient management. Of greater concern, the term 'deleterious' is an output from multiple tools (CONDEL, LRT, Mutation Taster, Provean); this term is also used by the European Medicines Agency (http://www.ema.europa.eu/) and the US Food and Drug administration (https://www.fda.gov/) to denote eligibility of patients with specified cancer types/presentation for poly ADP ribose polymerase inhibitor therapy, namely patients with a 'deleterious or suspected deleterious germline (or somatic) BRCA mutation'. Furthermore, the combined term 'deleterious mutation' is used (in addition to the term 'pathogenic mutation') by the NCCN 2018 guidelines (www.nccn.org) to describe genetic variation used to denote specific management recommendations for patients with familial breast-ovarian cancer. Without clarity of the use of these terms in context, there is significant risk of overinterpretation of bioinformatic data. Cancer genetic germline tests are increasingly being ordered by clinicians relatively unskilled in genetic terminology. A clear reporting language, with clear definitions of final variant interpretation summarising all the component information used for classification, is thus paramount to avoid variant misinterpretation and inappropriate patient management.

Proposed vocabulary to describe genetic variation in cancer predisposition genes

The terms discussed below primarily focus on describing germline variation in cancer genes, detected by genetic testing for diagnosis of hereditary cancer or estimating future cancer risk. However, the vocabulary inevitably overlaps terms used to describe somatic variation in tumours in the context of drug therapy selection for patients with cancer, or distinguishing true germline variants from variants arising from somatic

Table 1 Text descriptors from selected bioinformatic prediction programmes used for variant annotation in sequencing pipelines*

Programme	Output terms and other descriptions
CONsensus DELeteriousness score of missense mutations (CONDEL)	*Deleterious* Neutral (http://bg.upf.edu/fannsdb/) Description: the scores of different methods (SIFT, Polyphen2, Mutation Assessor, FATHMM, Ensembl-variation) are weighted using the complementary cumulative distributions of approximately 20 000 missense SNPs, both *deleterious* and neutral.
Functional Analysis through Hidden Markov Models (FATHMM)	Damaging Tolerated (http://fathmm.biocompute.org.uk)
Likelihood Ratio Test (LRT)	*Deleterious* Disease-causing (identified in a 'mutation database') Polymorphism (predicted OR annotated) (http://www.genetics.wustl.edu/jflab/lrt_query.html)
Mutation taster	*Deleterious* Neutral Unknown (http://www.mutationtaster.org)
Mutation assessor	Predicted non-functional (low, neutral) Predicted functional (low, neutral) (http://mutationassessor.org)
Predicting disease-associated non-synonymous SNPs Analyzer (nsSNPAnalyzer)	Disease Neutral (http://snpanalyzer.uthsc.edu/)
Predictor of human deleterious SNPs (PhD-SNP)	Disease Neutral (http://snps.biofold.org/phd-snp/)
Polymorphism Phenotyping v2 (PolyPhen-2)	Probably damaging Possibly damaging *Benign* (http://genetics.bwh.harvard.edu/pph2)
Protein Variation Effect Analyzer (Provean)	*Deleterious* Neutral (http://provean.jcvi.org)
Sorting Intolerant From Tolerant (SIFT)	Damaging Tolerated (http://sift-dna.org)

*Prediction tools used for missense variants denoted in bold are included as options for scoring bioinformatic predictions in the ClinGen Pathogenicity calculator[40] (http://calculator.clinicalgenome.org/site/cg-calculator), and the ClinGen Variant Curation Interface (https://curation.clinicalgenome.org/), both developed to enable application of the ACMG/AMP guidelines. The meta-predictors Rare Exome Variant Ensemble Learner (REVEL),[41] Combined Annotation Dependent Depletion (CADD)[42] and BayesDel,[43] provide continuous scores and not specific terms as output. Output terms also used in clinical reporting, or to define eligibility for poly ADP ribose polymerase inhibitor treatment, are noted in italics.
ACMG, American College of Medical Genetics and Genomics; AMP, Association for Molecular Pathology.

clonal drift in 'disease free' tissue used for DNA extraction.[13 14] These suggestions take into consideration terms put forward by the IARC unclassified sequence variants working group,[6] ACMG/AMP[5] and HVP,[15] and a comprehensive review article assessing clinical implications of gene panel test results for breast cancer risk prediction.[16] We have not addressed variant annotation in relation to predicting response to drug treatment. We refer readers to the Clinical Pharmacogenetics Implementation Consortium for consensus terms for reporting clinical pharmacogenetic results,[17] and note that ClinVar currently supports the following terms describing variant effect relating to therapy: drug response, confers sensitivity. For any given variant, the term *wild-type* may be used to denote the nucleotide/s or amino acids in the selected reference DNA/protein

sequence. However, this term can also be used to describe 'normal' phenotype, typically protein function/characteristics measured by in vitro assays.

The term *variant* should be used to define a DNA change that differs from a defined reference sequence, consistent with recommendations from the ACMG/AMP[5] and HVP.[15] Various descriptors of a variant depend on the context, as denoted below.

Cellular origin of variant

It is important to specify the tissue from which tested DNA has been derived, irrespective of the use of the descriptors below.

► *Constitutional or germline* (used interchangeably): a sequence variant identified in DNA from a tissue type *assumed* to represent the DNA content of the fused germ cells (eg, blood), and therefore to be transmittable to offspring. This includes a sequence variant that arises de novo in a gamete and in this setting will be present in all cells of an individual but not inherited from one or other parent.

► *Mosaic:* sequence variant that has arisen during embryogenesis and therefore not present in all the cells/tissues of an individual.

► *Somatically acquired* (not inherited): sequence variant present only in a specific tissue. In the context of tumour DNA (tumour biopsy or circulating tumour DNA derived from blood), the variant will be present in tumour DNA and absent from DNA derived from other tissue/s of the same individual.

► *Somatically detected:* sequence variant detected in a specific tissue type and for which somatic or germline origin has not yet been established by investigating DNA from other tissues. May be used for variation detected by tumour sequencing (*tumour-detected*), or in the context of suspected mosaicism. Somatically detected variants identified in DNA from blood/saliva with allele proportion <0.3, and/or in individuals with incompatible clinical presentation, are more likely to represent variation due to aberrant clonal expansion in hematopoietic cells (particularly *TP53*[13 14]), or from circulating tumour DNA.

Nucleotide-level evolutionary conservation

Nucleotide sequence changes in coding regions are primarily assessed using protein-level conservation analysis that assesses their effect on protein sequence (see below). However, nucleotide-level conservation analysis may be considered useful for investigating effect of sequence changes on the fitness of splicing regulatory motifs, or mRNA secondary structure and stability, translation efficiency[18–20] or to infer functional importance of non-coding sequences (introns, untranslated regions and other extragenic sequence). Indeed, it is a factor denoted for review of synonymous variants (code BP7) in the ACMG/AMP guidelines.[5]

Nucleotide substitutions analysed by evolutionary/phylogenetic methods involve alignment of at least three nucleic acid sequences, termed multiple (multispecies) sequence alignment (MSA). We suggest that such analysis specify the method/programme used, the number of ortholog sequences included and their phylogenetic relationship to humans. To our knowledge, there are no firm standards proposed for use of nucleotide-level evolutionary conservation in predicting whether a variant may affect fitness of difference sequence motifs (splicing, transcription factor binding, etc).

We thus suggest that nucleotide positions in the alignment may be described simply as:

Table 2 Alternative terms currently in use to describe five-tier disease gene variant classification categories

	IARC classification scheme[6] Intended use, highly penetrant cancer susceptibility genes				ACMG/AMP[5] Intended use, Mendelian diseases		
Numerical class	Terms	Probability of being pathogenic	Suggested predictive testing of at-risk relatives	Suggested surveillance*	Terms	Probability of being pathogenic†	Description of clinical relevance
5	Definitely pathogenic‡	>0.999	Yes	Full high-risk surveillance guidelines (for variant carriers)	Pathogenic	>0.999	Variant classified as pathogenic using the proposed classification scheme has met criteria informed by empirical data such that a healthcare provider can use the molecular testing information in clinical decision making.
4	Likely pathogenic	0.950–0.999	Yes	Full high-risk surveillance guidelines (for variant carriers)	Likely pathogenic	0.900–0.999	Sufficient evidence that a healthcare provider can use the molecular testing information in clinical decision making when combined with other evidence of the disease in question.
3	Uncertain§	0.050–0.949	No (recommend *research* testing of family members)	Based on family history and other risk factors	Uncertain significance	0.100–0.899	Should not be used in clinical decision making; efforts to resolve the classification of the variant as pathogenic or benign should be undertaken.
2	Likely not pathogenic or of *little* clinical significance¶	0.001–0.049	No (recommend *research* testing of family members)	Treat as 'no pathogenic variant detected' for this disorder (ie, based on family history and other risk factors)	Likely benign	0.001–0.099	Sufficient evidence that a healthcare provider can conclude that it is not the cause of the patient's disorder when combined with other information.
1	Not pathogenic or of *no* clinical significance**	<0.001	No	Treat as 'no pathogenic variant detected' for this disorder (ie, based on family history and other risk factors)	Benign	<0.001	Sufficient evidence that a healthcare provider can conclude that it is not the cause of the patient's disorder.

*Represented with minor modifications for clarity (words in parentheses) introduced by the ENIGMA consortium.

†ACMG/AMP guidelines do not require quantitative variant classification methods to be used, but nevertheless propose probabilities of a variant either being disease-causing or benign.[5]

‡Represented as 'pathogenic' by InSiGHT, ENIGMA and on the BRCA-Exchange website (http://brcaexchange.org/).

§Represented as 'uncertain significance' on the BRCA-Exchange website (http://brcaexchange.org/).

¶Represented as 'likely benign' on the BRCA-Exchange website (http://brcaexchange.org/).

**Represented as 'benign/little clinical significance' on the BRCA-Exchange website (http://brcaexchange.org/).

ACMG, American College of Medical Genetics and Genomics; AMP, Association for Molecular Pathology; ENIGMA, Evidence-Based Network for the Interpretation of Germline Mutant Alleles; IARC, International Agency for Research on Cancer; InSiGHT, International Society for Gastrointestinal Hereditary Tumours.

► *Evolutionarily invariant*: at the position of the variant, the MSA is identical across all species considered in the alignment.

► *Evolutionarily variant*: at the position of the variant, the MSA is not identical across all species considered.

Scores provided by specific tools, eg, PhyloP,[21] may be helpful to assess if a specific position is evolutionary constrained or not.[22] Furthermore, position weight matrices[23] developed for functionally important sequence motifs, eg, splice junctions[24] may be useful to gauge the effect of a genetic variant on the fitness of that sequence motif.

Protein-level evolutionary conservation and bioinformatically predicted physicochemical characteristics of a missense alteration

As noted above (table 1), bioinformatic tools use a range of terms to describe results from analysis of a given predicted missense alteration. Protein-level conservation analysis is required to adequately capture redundancy in codon usage, and additional features considered include relative physicochemical properties of amino acids, and predicted effects on protein secondary, tertiary and quaternary structure. Without prescribing or

recommending use of any particular tool/s for variant evaluation, we do recommend use of the following terms to describe output for analysis of missense substitutions (or small in-frame insertions/deletions) using evolutionary/phylogenetic methods. Depth of the analysis for a protein sequence alignment should be specified, including number of ortholog sequences in the protein multiple sequence alignment (PMSA), phylogenetic relationship of the species most evolutionarily distant to humans and the average number of substitutions per position.[25]

Variants should be described in relation to the level of evolutionary conservation for that amino acid position (residue) in the protein multiple sequence alignment (and noting that the non-human sequences included in the alignment should be wild-type (the form that occurs most frequently) and of a splice form matching the human reference sequence, insofar as possible).

Generic descriptors for an amino acid position (residue) in an alignment

► *Evolutionarily invariant:* amino acid at that position in the PMSA is identical across all species considered.

Table 3 Recommended terminology and descriptors for five-tier disease gene variant classification categories, considering variant pathogenicity in the multigene panel testing arena*

Numeric class	Consolidated five-tier description	Suggested acronym	Generic description of cancer risk determined for the variant	Generic description of relevance to clinical management for germline variants in cancer susceptibility genes	
				High penetrant variants (>fourfold risk relative to population)	Moderate penetrant variants (twofold to fourfold relative risk)
5	Pathogenic	P	Sequence variant is associated with ≥twofold cancer risk, and could be used to inform medical management.	Sequence variant may be used alone to inform clinical management. Management recommendations for an *individual* should be determined in accordance with absolute risk of specified cancer types, considering clinical presentation and other known genetic and environment risk factors.‡ Offer predictive testing for relatives.	Clinical management recommendations for a variant carrier should consider knowledge of personal and family history of disease, and other known genetic and environmental risk factors, that together can strongly influence absolute risk for an individual.‡ Consider predictive genetic testing for relatives only if supported by local (regional/national) guidance
4	Likely pathogenic	LP	Sequence variant is likely associated with ≥twofold cancer risk, and could be used to inform medical management.†	As above—sequence variant may be used alone to inform clinical management. Consider other factors to refine estimate of absolute risk for an individual.	As above—sequence variant should be used to inform clinical management only after consideration of other factors with influence absolute risk for an individual.‡
3	Uncertain (significance)	VUS	Sequence variant has been assessed for association with cancer phenotype/s but risk association remains uncertain.	Clinical management recommendations should be determined on the basis of personal and/or family history of disease, and other known genetic and environmental risk factors.‡ The presence of the variant should not be used to influence management of the carrier individual or their relatives. Research testing of family members may be recommended to aid variant classification.	
2	Likely little clinical significance/likely benign§	LB	Sequence variant is likely NOT associated with ≥twofold cancer risk.	Variant on its own is likely to be of no or little clinical significance. Clinical management recommendations should be determined on the basis of personal and/or family history of disease, and other known genetic and environmental risk factors.‡ Further research may clarify variant contribution (if any) to risk.	
1	Little clinical significance/benign§	B	Sequence variant is NOT associated with ≥twofold cancer risk.	Variant on its own is of no or little clinical significance. Clinical management recommendations should be determined on the basis of personal and/or family history of disease and other known genetic and environmental risk factors.‡	

*The tier descriptions have been adapted to: allow for both high-risk and moderate-risk variants (irrespective of the gene involved) to be annotated for medical actionability in accordance with the level of risk/s they impart to individual carriers; consider that relative risks are age-specific for common diseases such as cancer where incidence in the general population increases with increasing age, so the relative risk associated with a cancer predisposition gene falls with increasing age; denote specifically that clinical management recommendations should consider personal and family history of disease, as well as environmental exposures, and other genetic risk factors (in particular polygenic risk score information). *Terminology assumes that only variants associated with a relative risk of >twofold will be reported out as unique variants with directly assigned pathogenicity.*

†Defined as 90% (ACMG/AMP) or 95% (IARC) certainty of being pathogenic or benign. As per IARC recommendations,[6] further research, including research testing of family members, may be helpful to better determine the risk association and clinical significance of the variant.

‡Other factors may reduce or increase the risk of disease. Risk factors to be assessed may include polygenic risk scores, which themselves include information about individual variants associated with <twofold risk (low increased risk). *Note*: inclusion of both family history and polygenic risk score information for absolute risk estimation should account for the proportion of familial relative risk that is explained by genetic factors included in the polygenic risk score calculation.[44] Further implementation research is required to understand how best to implement PRS testing to stratify cancer risks in a range of settings, including patients with cancer and general population screening.

§The combined text description was selected as preferable for *initial* presentation in reporting, to underscore the fact that some sequence variants falling into class 1 or class 2 may be causally associated with a defined low increased risk of cancer, eg, the *BRCA2* c.9976A>T p.(Lys3326Ter) variant associated with <1.5-fold increased risk of breast or ovarian cancer.[37]

► *Evolutionarily conserved:* amino acids at that position in the PMSA have *similar** physicochemical properties across all species considered.

► *Not evolutionarily conserved:* amino acids at that position in the PMSA show *marked differences** in physicochemical properties across the species considered.

*There are alternative methods to assess similarity and differences for substitutions at a given position in an MSA. The method should be defined for the specific analysis conducted. Examples include: Grantham variation (GV) is <60 (conserved) or ≥60 (not conserved); residue harbours an alternate amino acid with Grantham difference (GD) score <60 (conserved) or residue variation exceeds this limit (not conserved).[26]

Descriptors for an amino acid change relative to the sequence alignment

► *Outside the range of variation (observed evolutionarily):* altered amino acid has markedly different physicochemical properties (defined by size, charge, etc) to the range of variation of those properties observed at its position in the PMSA. *Note*: this is relatively more likely to happen if the position is invariant or conserved.

Table 4 Suggested approach to multitier reporting of cancer gene variants conferring high or moderate disease risk*

Demographic information	
Cancer phenotype	<specify cancer phenotype of proband>
Sample tissue	□ Blood □ Saliva □ Primary tumour □ Distant metastasis □ Other (specify)
Context	□ Diagnostic □ Prediction cancer risk □ Genotype directed treatment □ Other (specify)
Variant identified	Variant description should be based on the Human Genome Organisation (HUGO) standard variant nomenclature and Human Genome Variation Society (HGVS). The use of HGVS nomenclature can be problematic for describing exon deletions/duplications (particularly where end points are unknown) and triplet repeat expansions. Therefore, such variants should also be described in words if this improves clarity.
Level 1: variant classification*	**Variant interpretation by the laboratory—variant classification based on all available data.**
ACMG/AMP or IARC variant classification	Pathogenic *or* likely pathogenic *or* uncertain significance *or* likely little clinical significance/likely benign *or* little clinical significance/benign. Assertion relates to dominant mode of inheritance for cancer phenotype. Details of the evidence supporting the variant classification are provided as supplementary documentation.
Level 2: clinical validity*	**Considers the strength of evidence for the genetic variant being related to the presenting cancer phenotype, new primary cancer risk, predicting the likely response to targeted treatment or relevance to recessive phenotypes. It is recognised that specific missense and protein truncating variants within the same gene may exhibit a differing magnitude of effect on cancer risk.**
Presenting cancer phenotype	There is strong evidence that this variant is making a (*substantial*)*contribution* to the presenting cancer phenotype <insert reference>. *or* There is insufficient evidence to support an association between this variant and the presenting cancer phenotype <insert reference if available>.†
Prediction cancer risk	There is strong evidence to support the prediction of a *high* (>fourfold) increase in future cancer risk <specify cancer type/s> when this variant is present in a family member *(reference with associated risks)*. *or* There is strong evidence to support a *moderate* (twofold to fourfold) increase in future cancer risk <specify cancer type/s> when this variant is present in an family member and can be used for cancer risk stratification. This variant should not be considered in isolation. Information from the family history of disease, and other known genetic and environmental risk factors or polygenic risk scores may substantially modify overall cancer risk estimates <insert reference with associated risks>. *or* There is strong evidence from population studies to support that is variant is associated with a low (<twofold) increase in cancer risk <specify cancer type/s><reference with associated risks>. This variant is insufficiently predictive of future risk to be clinically actionable. Variants in this category may contribute towards a polygenic risk score.
Genotype directed treatment	There is evidence to support consideration of <specify drug/drug type> in the context of <specify cancer type><insert reference/s>.‡ *or* There is currently <limited/no> evidence to support consideration of <specify drug/drugtype> in the context of <specify cancer type>.
Biallelic inheritance	Evidence may support that biallelic (compound heterozygote or homozygote) variant inheritance is likely to cause recessive disease<specify disease name>.
Level 3: clinical utility and actionability*	**This final element comprises the discussion between physician and patient. It may take the form of a personalised assessment of risk based on the presenting cancer phenotype, clinical scenario and family history. If the reporting laboratory is not qualified to address this element of the report then it should be made clear that this is an additional requirement before determining any clinical management consequences. Proposed clinical interventions should be *risk proportionate* and take the *individual clinical circumstances* into account reflecting on any uncertainty around estimates of risk underpinning life-changing decisions such as risk reducing surgery or reproductive choices. It also requires consideration of cascade genetic testing for other relatives at risk. If preferred, the report may be shortened by referring to local guidelines for details.**
(i) Clinically actionable (high penetrance)	*Simplified report:* follow clinical management guidelines for high penetrance predisposition genes according to local guidelines *Or* *Detailed report* presenting details from local guidelines—EXAMPLE provided: 1. *Surveillance:* high-risk surveillance if strong evidence for variant-specific high risk 2. *Risk reducing surgery:* consider risk reducing surgery *only* if the overall clinical picture is high risk (see above) and depending on cancer prognosis and treatment <specify appropriate risk reducing surgery>. 3. *Cascade genetic testing:* sequence variant may be used alone to inform clinical management and so cascade genetic testing is indicated.
(ii) Clinically actionable but not in isolation (moderate penetrance)	*Simplified report:* manage based on a comprehensive risk evaluation§ Additional moderate-risk or high-risk surveillance may be indicated. Clinical management recommendations should be determined on the basis of the absolute cancer risks conferred by <variant identified> in combination with the personal and/or family history of disease and other known genetic and environmental risk factors. Follow clinical management guidelines according to local guidelines. *Or* *Detailed report* presenting details from local guidelines—EXAMPLE provided: 1. Manage based on a comprehensive risk evaluation‡: additional moderate-risk or high-risk surveillance may be indicated. Clinical management recommendations should be determined on the basis of the absolute cancer risks conferred by <variant identified> in combination with the personal and/or family history of disease and other known genetic and environmental risk factors. 2. Risk reducing surgery: for moderate penetrance gene variants in isolation, there is currently no clear evidence of clinical benefit for risk reducing surgery. 3. Cascade genetic testing: predictive testing for this variant has limited clinical utility in isolation.
(iii) Not clinically actionable (low penetrance)	Manage based on family history: Insufficiently predictive of future cancer risk to be clinically actionable. Clinical management recommendations for the <presenting cancer phenotype> should be determined on the basis of personal and/or family history of disease and other known genetic and environmental risk factors. Variants in this category may contribute towards a polygenic risk score.

*We suggest that this should be repeated for each reportable variant identified in the sample submitted (usually the proband—defined as the person serving as the starting point for the genetic study of a family; may be a patient with cancer or not). For high-risk cancer susceptibility genes, the probability threshold for classification as likely pathogenic is 0.95 for the IARC classification scheme[6] and 0.90 for the ACMG/AMP guidelines.[5] It may be reasonable to consider the 0.90 threshold as more appropriate for moderate penetrance variants, where recommended management excludes irreversible surgical risk-reducing strategies. We suggest that only strong evidence supporting risk associations should be used to determine clinical validity of clinically actionable variants. We define strong evidence following recommendations published in the study by Easton *et al*,[16] namely: "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".

†We suggest that individual results for risk alleles associated with small increase in risk of cancer (as determined by well-powered studies) should not be included in clinical genetic test reports in isolation but presented as a combined overall risk prediction score.

‡We do note that it cannot be assumed that all variants that are associated with increased disease risk will predict (the same) response to targeted therapy and vice versa. It is thus recommended that future iterations of multitier reporting schemes provide for distinct annotation of germline variants for disease risk and relevance to drug treatment.

§High-risk surveillance if comprehensive cancer risk (family history-based risk) stratification >30% absolute lifetime risk, moderate-risk surveillance if comprehensive cancer risk stratification (family history-based risk) 17%–30% absolute lifetime risk, population screening if comprehensive cancer risk stratification (family history-based risk) <17% absolute lifetime risk.

ACMG, American College of Medical Genetics and Genomics; AMP, Association for Molecular Pathology; IARC, International Agency for Research on Cancer.

► *Similar to the range of variation (observed evolutionarily):* altered amino acid has similar physicochemical properties to the extremes observed for the range of variation of physicochemical properties at that position in the PMSA, for example, GV>0 and GD relatively small, say <30.

► *Inside the range of variation (observed evolutionarily):* altered amino acid has physicochemical properties that clearly fall within the range of variation of those physicochemical properties observed at that position in the PMSA, for example, GV>0 and GD=0.

If the position of an amino acid variant in the PMSA is invariant or conserved, and the change is outside the range of variation, then it is considered *evolutionarily unlikely*. Conversely, an amino acid substitution that is within or similar to the range for variation observed evolutionarily, may be termed *evolutionarily tolerated* (if the alternative amino acid is already present in the alignment) or otherwise *evolutionarily tolerable* (if the alternative amino acid is not observed in the alignment, but similar to the range of variation observed).

As noted above, bioinformatic prediction of variant effect on function should not be used alone to infer association with measurable disease risk. However, variant effect/bioinformatic prediction scores, together with information on variant location in the gene relative to splicing motifs/functional domains, may be calibrated against clinical measures of variant pathogenicity (termed *clinical calibration*) to provide probability estimates useful to re-assign a variant as likely not pathogenic[27–30] (see online supplementary text for more details).

Impact on mRNA transcript profile or protein function

mRNA profile

We recommend 'naturally occurring mRNA transcript' be used to describe mature mRNA transcript/s seen in controls. Using mRNA transcription in control samples as reference, a variant may exhibit an altered mRNA transcript profile by: (i) impacting overall level of transcript/s (overall expression); (ii) resulting in novel mature mRNA transcript/s and/or (iii) altering the relative contribution of individual transcripts to the overall expression. Control mRNA should be from the same tissue type and analysed using the same methodology.

Variants assessed for effect on *transcription* via gene regulation, may be described as not impacting transcription levels, or impacting transcription levels. Impact on transcription can be further described as partial, or total (also termed transcriptional silencing). Epigenetic silencing specifically refers to impact on transcription via altered methylation profile.

Variants assessed for effect on mRNA transcript profiles via impact on mRNA splicing, including loss, gain or enhanced use of cryptic splicing motifs, may be described as follows:

► *Non-spliceogenic:* the variant does not alter mRNA transcript profile.

► *Spliceogenic (predicted) LOF:* the variant results in an altered mRNA transcript profile that is predicted to cause gene loss-of-function, that is, any combination of mRNA transcripts predicted non-coding, predicted protein truncating nonsense mediated decay (NMD) and/or predicted to encode proteins lacking critical structural/functional motifs.

► *Spliceogenic (predicted) functional:* the variant results in an altered mRNA transcript profile that is predicted to preserve gene functionality, that is, any combination of mRNA transcripts which together will encode protein/s that is/are predicted to preserve functional capacity.

► *Spliceogenic uncertain function:* the variant results in an altered mRNA transcript profile for which the *coding/functional consequences are uncertain*, that is, combinations of transcripts predicted to cause gene loss-of-function, retain gene function or to encode proteins with uncertain functional potential, for which the combined functional capacity is unclear.

Protein function

Variants that have been analysed in functional (biochemical, biophysical, molecular biological) assays that assess variant effect on protein conformation/activity/function should compare effect (always specifying effect measured) to wild-type and other controls as follows:

► *No functional impact:* variant displays features (*specified*) similar to wild-type.

► *Functional impact:* variant alters features (*specified*) compared with wild-type. Impact may be described as:
  – *Complete loss of function:* variants with loss of function (feature to be specified) below a detection threshold or to a degree of the average pathogenic variant for that gene/protein.
  – *Partial loss of function:* variants with partial loss of function (feature to be specified), that is, intermediate between that of the wild-type protein sequence and the average pathogenic variant for that gene/protein. May alternatively be described as *intermediate functional effect* or *hypomorphic*.
  – *Gain-of-function:* term encompasses increase in a known function for that protein relative to wild-type, *or* gain of additional novel functions, for example, for p53,[31] RET.[32] May alternatively be described as *neomorphic*.
  – *Dominant-negative:* variant that encodes an altered protein that interferes with the function of the protein encoded by the wild-type allele. A common example is a variant encoding a protein that retains the ability to form protein-protein complexes, but disrupts the functionality of such complexes.

Note: a variant with measurable effect in vitro on mRNA transcript profile or protein function (specifying feature measured), relative to appropriate controls, should not a priori be assumed to be associated with disease risk. To include functional and mRNA data in gene-specific variant classification protocols, it is necessary that the association between magnitude of effect on mRNA profile/protein function and disease risk is first calibrated against clinical measures of variant pathogenicity, such that the range of variation in effect is established for variants previously classified as pathogenic, and for those considered not pathogenic. See de la Hoya *et al* and Colombo *et al*[33 34] for examples of calibration of *BRCA1* and *BRCA2* transcript levels.

Genetic variation and description of associated disease risk

Cancer risks associated with a genetic variant may be presented in a variety of different ways. Risk associated with a proven cancer-predisposing gene variant (type) can only be correctly interpreted if the time period and population to which the risk applies is defined.[35] Most cancer predisposition genes exhibit organ-specific *disease expressivity*, so it is important to specify disease (phenotype), and mode of inheritance. A given variant may confer different disease risks for heterozygote versus compound heterozygote or homozygote carriers.

► *Absolute or cumulative risk* is the likelihood that a person with a cancer-predisposing variant will develop a given

cancer within a period of time, for example, within the next 5 or 10 years, or by a specific age. It is expressed as a percentage.

► *Relative risk* compares the cancer risk for genetic variant carriers relative to the risk for non-carriers or the general population, and can be estimated through several study designs, for example, case-control studies estimate odds ratios, cohort studies estimate rate ratios.

► *Disease penetrance* is typically used to describe the overall probability that carriers of cancer-predisposing variants in a given gene (sometimes specifying a specific variant type) will develop specified cancer type/s until a specified age or during lifetime. For a fully penetrant genetic variant (or variant type), disease will develop in all individuals with the variant (type). *Reduced penetrance* may be used to describe a variant that displays lower penetrance compared with risk-associated variants typically identified for that disease gene. The estimated level and type of disease risk/s associated with a reduced penetrant variant determine whether carrier status may be used to inform clinical management.

We suggest that it is helpful to present variant-associated risks to patients as both an absolute measure (eg, 50 in every 100 people with this variant (type) are expected to develop breast or ovarian cancer by age 70 years) and a relative measure (eg, a variant carrier is 10 times more likely to develop breast cancer in their lifetime compared with women in the general population), and report these with appropriate confidence intervals (). Based on descriptors applied previously for breast cancer,[16] for this discussion document we have categorised cancer risk levels associated with a given variant, relative to the general population risk, as follows: high increased risk, more than fourfold; moderate increased risk, twofold to fourfold; low increased risk, greater than unity and less than twofold. Relative risks are not clinically useful without knowing the absolute risk of a disease—a relative risk of four for a rare disease is still a small risk. A high *relative* risk is not necessarily a high *absolute* risk because the latter depends on the baseline population risk. Thus, for cancer types that are uncommon in the population, the absolute risk, and also the availability of interventions, have to be considered when determining the clinical actionability of a variant. *Note*: the term 'intermediate' requires reference values to define its level (for relative or absolute risk), and is thus considered non-specific for the purpose of variant reporting.

The term *risk allele* may be used as an alternative to describe a variant identified as cancer-associated, generally using case-control analysis such as genome-wide association studies, where there is not necessarily a mechanistic relationship between a 'lead' variant in a linkage disequilibrium block and disease predisposition.

Proposed vocabulary to describe clinical relevance of genetic variation in known or suspected cancer predisposition genes using a five-tier system

The IARC five-tier variant classification system was developed to promote use of probability-based methods for variant classification of highly penetrant cancer susceptibility genes that could then be specifically linked to recommended clinical management protocols.[6] This system has been adopted by the InSiGHT group for mismatch repair (MMR) gene variant classification,[36] and by ENIGMA for *BRCA1/2* variant classification (https://enigmaconsortium.org). It is used for ClinGen-approved expert panel curation of variants in these genes, displayed in ClinVar (https://www.ncbi.nlm.nih.gov/clinvar/) and selected public

locus-specific databases (https://www.insight-group.org/variants/; http://brcaexchange.org/). The IARC tier terminology and management recommendations as published in 2008 are broadly consistent with those recommended by ACMG/AMP (table 2). However, assigning terms for the variant tiers across different public portals has highlighted differences in the wording used to describe the IARC class 2 and class 1 tiers, and potential for misinterpreting the clinical relevance of individual variants based on current IARC or ACMG/AMP terms. Indeed, misinterpretation of the class 1 tier has been raised in relation to the *BRCA2* c.9976A>T p.Lys3326Ter variant associated with <1.5-fold increased risk of breast or ovarian cancer,[37] both publicly,[38 39] and by direct query to the BRCA-Exchange website (http://brca-exchange.org/). The latter led to a change in representation of this tier as 'benign' to 'benign-little clinical significance' on the BRCA-Exchange website.

Furthermore, during development of the ENIGMA *BRCA1/2* variant classification criteria (https://enigmaconsortium.org), research results emphasised the need for clear statements about appropriate class assignment for variants with proven association with so-called 'intermediate' or 'moderate' increased risk of cancer. Specifically, discovery that the *BRCA1* c.5096G>A p.Arg1699Gln variant demonstrates reduced disease penetrance relative to 'high-risk' truncating *BRCA1* variants raised the issue of how to denote such reduced penetrance variants in the five-tier system, in particular if the disease penetrance was sufficient to trigger altered management, although not as extensive as the 'standard pathogenic' variant for that gene.[3] The advent of multigene panel testing that encompasses so-called 'moderate-risk genes' has further highlighted the complexities of trying to develop and implement simple terms to describe the disease risk and clinical relevance of variants where risk by variant type can differ between and within genes. Indeed, circulation and discussion of the ENIGMA terminology highlighted 'pathogenic' as the term for which the definition was most contentious.

In an attempt to address all the above issues, we considered usability of terms in research publications, inconsistencies in wording for the IARC class 1 and 2[6] and alignment with terminology recommended by the ACMG/AMP guidelines.[5] We also considered relevant definitions from several English dictionaries, and the derivation of the word (see online supplementary text)—this being an important component of translating meaning of terms by collaborators for whom English is not the first language.

During the ENIGMA meetings held on January 2017, September 2017 and June 2018, the ENIGMA membership have been presented with various options for describing or rewording terms, with more detailed descriptions of each of the five tiers intended to capture the complexity of reporting in the multigene panel testing era. Discussions arising from these presentations, and additional commentary on documentation circulated to members, has resulted in the recommendations and summary descriptions shown in table 3. We anticipate that this more detailed description of the clinical implications of, and management recommendations associated with, germline variants placed in each of the classification tiers will provide a short-term solution to improve understanding of these terms in the context of clinical reporting of cancer predisposition variants using a five-tier classification system. Adaptation for other Mendelian or co-dominant disease genes is possible, subject to clear definition of level of disease risk associated with clinical actionability, and other factors to be considered when establishing absolute risk at the individual level.

> **Box 1 Key recommendations regarding variant descriptors and their use in variant classification and clinical reporting***
>
> ► The term *variant* should be used to define a DNA change that differs from a defined reference sequence.
> ► It is important to always specify the tissue from which tested DNA has been derived.
> ► Bioinformatic prediction of variant effect on function should not be used alone to infer association with measurable disease risk.
> ► Bioinformatic prediction scores, *together with* information on variant location in the gene relative to splicing motifs/functional domains, may be calibrated against clinical measures of variant pathogenicity (termed *clinical calibration*) to provide probability estimates useful to re-assign a variant as likely not pathogenic.
> ► The term *spliceogenic* is used generically to describe a variant that results in altered mRNA transcript profile (relative to a reference), without consideration of transcript/s susceptibility to NMD, ability to encode functional protein or association with disease risk.
> ► Variants analysed in *functional assays* (biochemical, biophysical, molecular biological, cellular) that assess variant effect on protein conformation/stability/activity/function should describe effect compared with wild-type and other controls, and always specify the protein effect measured.
> ► A variant with measurable effect in vitro on mRNA transcript profile or protein function (specifying feature measured), relative to appropriate controls, should not a priori be assumed to be associated with quantifiable disease risk.
> ► It is critical to specify disease/phenotype and mode of inheritance when providing a pathogenicity assertion for a genetic variant.
> ► Present variant-associated risks as an absolute measure, and a relative measure, and report these with appropriate CIs.
>
> *Variant annotation is a broad term used in the context of next-generation sequencing bioinformatic pipelines to describe the process of assigning a variety of descriptors to a given sequence variant, but these annotations are largely distinct from the clinically focused variant terminology denoted above (see online supplementary text for an overview of variant annotation).

Proposal for development of a multitier system for variant annotation in clinical test reporting of multigene panel results

Despite the expansion of descriptions for the five-tier variant classification system shown in table 3, it was clear from comments received that assignment of variant pathogenicity using the current five-tier system is inadequate to deal with the complexities of reporting multigene panel testing outcomes, and to portray differences in variant-specific risks for a given gene. The term 'pathogenic' remained contentious, with comments raised by ENIGMA members including: need to capture the relevance of genetic findings to patient disease diagnosis (phenotype) versus relevance of a secondary finding (ie, outside of the patient diagnosis); reporting variant effect for recessive as well as dominant disease and whether a variant could be termed 'pathogenic' on the background of a polygenic risk score that *reduced* individual risk to the population level. These observations indicate a need for a more consistent approach to variant reporting for clinical use, to minimise ambiguity of clinical management

considerations. We thus developed a template to emphasise the value of a multitier reporting system (outlined in table 4), and provide several worked examples (online supplementary table 2) to indicate its potential to capture the complexity of clinical actionability for variants identified by multigene cancer panel testing. The intention is that clinical inferences should be added to specific variant interpretation/classification, requiring the report to capture the level of (un)certainty around risk estimates and the contribution of an individual reported variant to a composite risk score. This could then be linked to clinical discussions about potential interventions, with particular value for multigene panel reporting.

CONCLUSIONS AND FUTURE DIRECTIONS

Our international consortium experience has highlighted that many terms used to describe genetic variants have multiple meanings, so that terms may be used interchangeably with the potential for false inferences in different contexts. Variant descriptor output from bioinformatics tools has potential to lead to patient mismanagement if directly transferred into clinical reports without clear explanation. Furthermore, there is considerable debate regarding use of terms to describe risk association and relevance to clinical management, with particular contention around the term pathogenic and relationship with patient medical management. We summarise the key points and provide recommendations on variant annotation and terminology in box 1. We also propose a framework for describing variants using a vocabulary that may be incorporated into clinical laboratory reporting. If adopted this approach should lead to more consistent variant interpretation at the laboratory level (ACMG/IARC), and importantly, allow clinical reports to clearly capture the relevance of a variant (or combination of variants) for the intended healthcare application. We recognise that practical implementation of such a system would require routine input from appropriately trained clinicians before a test report is issued for discussion with the patient. By no means intended as a final product, we present this for discussion and further development with the broader clinical community worldwide.

Author affiliations
[1]Genetics and Computational Biology, QIMR Berghofer Medical Research Institute, Herston, Queensland, Australia
[2]Faculty of Medicine, University of Southampton, Southampton, UK
[3]Centre for Cancer Genetic Epidemiology, University of Cambridge, Cambridge, UK
[4]School of Chemistry and Molecular Biosciences, University of Queensland, Brisbane, Queensland, Australia
[5]Medical Oncology Department, Hospital Clínico San Carlos, Madrid, Spain
[6]Basser Center for BRCA, Abramson Cancer Center, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, Pennsylvania, USA
[7]Gynaecology Research Unit, Hannover Medical School, Hannover, Germany
[8]Cambridge University Hospitals National Health Service Foundation Trust, Cambridge, UK
[9]Cancer Epidemiology, H. Lee Moffitt Cancer Center and Research Institute, Tampa, Florida, USA
[10]Department of Human Genetics, Radboud University Medical Center, Nijmegen, The Netherlands
[11]Department of Research, Fondazione IRCCS Istituto Nazionale dei Tumori, Milano, Italy
[12]Clinical Genetics Service, Memorial Sloan Kettering Cancer Center, New York City, New York, USA
[13]Department of Medical Genetics, Cambridge University, Cambridge, UK
[14]Division of Genetics and Epidemiology, Institute of Cancer Research, London, UK
[15]William Harvey Research Institute, Queen Mary Hospital, London, UK
[16]Human Genetics, Leiden University Medical Center, Leiden, The Netherlands
[17]Department of Pathology and Biomedical Science, University of Otago, Christchurch, New Zealand
[18]Oncological Sciences, University of Utah School of Medicine, Salt Lake City, Utah, USA
[19]Huntsman Cancer Institute, University of Utah, Salt Lake City, Utah, USA

REFERENCES

1 Spurdle AB, Couch FJ, Parsons MT, McGuffog L, Barrowdale D, Bolla MK, Wang Q, Healey S, Schmutzler R, Wappenschmidt B, Rhiem K, Hahnen E, Engel C, Meindl A, Ditsch N, Arnold N, Plendl H, Niederacher D, Sutter C, Wang-Gohrke S, Steinemann D, Preisler-Adams S, Kast K, Varon-Mateeva R, Ellis S, Frost D, Platte R, Perkins J, Evans DG, Izatt L, Eeles R, Adlard J, Davidson R, Cole T, Scuvera G, Manoukian S, Bonanni B, Mariette F, Fortuzzi S, Viel A, Pasini B, Papi L, Varesco L, Balleine R, Nathanson KL, Domchek SM, Offitt K, Jakubowska A, Lindor N, Thomassen M, Jensen UB, Rantala J, Borg Å, Andrulis IL, Miron A, Hansen TV, Caldes T, Neuhausen SL, Toland AE, Nevanlinna H, Montagna M, Garber J, Godwin AK, Osorio A, Factor RE, Terry MB, Rebbeck TR, Karlan BY, Southey M, Rashid MU, Tung N, Pharoah PD, Blows FM, Dunning AM, Provenzano E, Hall P, Czene K, Schmidt MK, Broeks A, Cornelissen S, Verhoef S, Fasching PA, Beckmann MW, Ekici AB, Slamon DJ, Bojesen SE, Nordestgaard BG, Nielsen SF, Flyger H, Chang-Claude J, Flesch-Janys D, Rudolph A, Seibold P, Aittomäki K, Muranen TA, Heikkilä P, Blomqvist C, Figueroa J, Chanock SJ, Brinton L, Lissowska J, Olson JE, Pankratz VS, John EM, Whittemore AS, West DW, Hamann U, Torres D, Ulmer HU, Rüdiger T, Devilee P, Tollenaar RA, Seynaeve C, Van Asperen CJ, Eccles DM, Tapper WJ, Durcan L, Jones L, Peto J, dos-Santos-Silva I, Fletcher O, Johnson N, Dwek M, Swann R, Bane AL, Glendon G, Mulligan AM, Giles GG, Milne RL, Baglietto L, McLean C, Carpenter J, Clarke C, Scott R, Brauch H, Brüning T, Ko YD, Cox A, Cross SS, Reed MW, Lubinski J, Jaworska-Bieniek K, Durda K, Gronwald J, Dörk T, Bogdanova N, Park-Simon TW, Hillemanns P, Haiman CA, Henderson BE, Schumacher F, Le Marchand L, Burwinkel B, Marme F, Surovy H, Yang R, Anton-Culver H, Ziogas A, Hooning MJ, Collée JM, Martens JW, Tilanus-Linthorst MM, Brenner H, Dieffenbach AK, Arndt V, Stegmaier C, Winqvist R, Pylkäs K, Jukkola-Vuorinen A, Grip M, Lindblom A, Margolin S, Joseph V, Robson M, Rau-Murthy R, González-Neira A, Arias JI, Zamora P, Benítez J, Mannermaa A, Kataja V, Kosma VM, Hartikainen JM, Peterlongo P, Zaffaroni D, Barile M, Capra F, Radice P, Teo SH, Easton DF, Antoniou AC, Chenevix-Trench G, Goldgar DE. ABCTB Investigators EMBRACE Group GENICA Network HEBON Group kConFab Investigators. Refined histopathological predictors of BRCA1 and BRCA2 mutation status: a large-scale analysis of breast cancer characteristics from the BCAC, CIMBA, and ENIGMA consortia. *Breast Cancer Res* 2014;16:3419.

2 Spurdle AB, Whiley PJ, Thompson B, Feng B, Healey S, Brown MA, Pettigrew C, Van Asperen CJ, Ausems MG, Kattentidt-Mouravieva AA, van den Ouweland AM, Lindblom A, Pigg MH, Schmutzler RK, Engel C, Meindl A, Caputo S, Sinilnikova OM, Lidereau R, Couch FJ, Guidugli L, Hansen T, Thomassen M, Eccles DM, Tucker K, Benitez J, Domchek SM, Toland AE, Van Rensburg EJ, Wappenschmidt B, Borg Å, Vreeswijk MP, Goldgar DE. kConFab Dutch Belgium UV Consortium German Consortium of Hereditary Breast and Ovarian Cancer French COVAR group collaborators ENIGMA Consortium. BRCA1 R1699Q variant displaying ambiguous functional abrogation confers intermediate breast and ovarian cancer risk. *J Med Genet* 2012;49:525–32.

3 Moghadasi S, Meeks HD, Vreeswijk MP, Janssen LA, Borg Å, Ehrencrona H, Paulsson-Karlsson Y, Wappenschmidt B, Engel C, Gehrig A, Arnold N, Hansen TVO, Thomassen M, Jensen UB, Kruse TA, Ejlertsen B, Gerdes AM, Pedersen IS, Caputo SM, Couch F, Hallberg EJ, van den Ouweland AM, Collée MJ, Teugels E, Adank MA, van der Luijt RB, Mensenkamp AR, Oosterwijk JC, Blok MJ, Janin N, Claes KB, Tucker K, Viassolo V, Toland AE, Eccles DE, Devilee P, Van Asperen CJ, Spurdle AB, Goldgar DE, García EG. The *BRCA1* c. 5096G>A p.Arg1699Gln (R1699Q) intermediate risk variant: breast and ovarian cancer risk estimation and recommendations for clinical management from the ENIGMA consortium. *J Med Genet* 2018;55:15–20.

4 Eccles DM, Mitchell G, Monteiro AN, Schmutzler R, Couch FJ, Spurdle AB, Gómez-García EB. ENIGMA Clinical Working Group. BRCA1 and BRCA2 genetic testing-pitfalls and recommendations for managing variants of uncertain clinical significance. *Ann Oncol* 2015;26:2057–65.

5 Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, Grody WW, Hegde M, Lyon E, Spector E, Voelkerding K, Rehm HL. ACMG Laboratory Quality Assurance Committee. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med* 2015;17:405–23.

6 Plon SE, Eccles DM, Easton D, Foulkes WD, Genuardi M, Greenblatt MS, Hogervorst FB, Hoogerbrugge N, Spurdle AB, Tavtigian SV. IARC Unclassified Genetic Variants Working Group. Sequence variant classification and reporting: recommendations for improving the interpretation of cancer susceptibility genetic test results. *Hum Mutat* 2008;29:1282–91.

7 MacArthur DG, Manolio TA, Dimmock DP, Rehm HL, Shendure J, Abecasis GR, Adams DR, Altman RB, Antonarakis SE, Ashley EA, Barrett JC, Biesecker LG, Conrad DF, Cooper GM, Cox NJ, Daly MJ, Gerstein MB, Goldstein DB, Hirschhorn JN, Leal SM, Pennacchio LA, Stamatoyannopoulos JA, Sunyaev SR, Valle D, Voight BF, Winckler W, Gunter C. Guidelines for investigating causality of sequence variants in human disease. *Nature* 2014;508:469–76.

8 Chenevix-Trench G, Milne RL, Antoniou AC, Couch FJ, Easton DF, Goldgar DE. CIMBA. An international initiative to identify genetic modifiers of cancer risk in BRCA1 and BRCA2 mutation carriers: the Consortium of Investigators of Modifiers of BRCA1 and BRCA2 (CIMBA). *Breast Cancer Res* 2007;9:104.

9 Griffith M, Spies NC, Krysiak K, McMichael JF, Coffman AC, Danos AM, Ainscough BJ, Ramirez CA, Rieke DT, Kujan L, Barnell EK, Wagner AH, Skidmore ZL, Wollam A, Liu CJ, Jones MR, Bilski RL, Lesurf R, Feng YY, Shah NM, Bonakdar M, Trani L, Matlock M, Ramu A, Campbell KM, Spies GC, Graubert AP, Gangavarapu K, Eldred JM, Larson DE, Walker JR, Good BM, Wu C, Su AI, Dienstmann R, Margolin AA, Tamborero D, Lopez-Bigas N, Jones SJ, Bose R, Spencer DH, Wartman LD, Wilson RK, Mardis ER, Griffith OL. CIViC is a community knowledgebase for expert crowdsourcing the clinical interpretation of variants in cancer. *Nat Genet* 2017;49:170–4.

10 Fokkema IF, Taschner PE, Schaafsma GC, Celli J, Laros JF, den Dunnen JT. LOVD v.2.0: the next generation in gene variant databases. *Hum Mutat* 2011;32:557–63.

11 Lindor NM, Guidugli L, Wang X, Vallée MP, Monteiro AN, Tavtigian S, Goldgar DE, Couch FJ. A review of a multifactorial probability-based model for classification of BRCA1 and BRCA2 variants of uncertain significance (VUS). *Hum Mutat* 2012;33:8–21.

12 Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS, Sunyaev SR. A method and server for predicting damaging missense mutations. *Nat Methods* 2010;7:248–9.

13 Coffee B, Cox HC, Kidd J, Sizemore S, Brown K, Manley S, Mancini-DiNardo D. Detection of somatic variants in peripheral blood lymphocytes using a next generation sequencing multigene pan cancer panel. *Cancer Genet* 2017;211:5–8.

14 Weitzel JN, Chao EC, Nehoray B, Van Tongeren LR, LaDuca H, Blazer KR, Slavin T, Facmg D, Pesaran T, Rybak C, Solomon I, Niell-Swiller M, Dolinsky JS, Castillo D, Elliott A, Gau CL, Speare V, Jasperson K. Somatic TP53 variants frequently confound germ-line testing results. *Genet Med* 2018;20.

15 Project) HHV. *HVP Guideline HVP/GL/003-01/EN*, 2016.

16 Easton DF, Pharoah PD, Antoniou AC, Tischkowitz M, Tavtigian SV, Nathanson KL, Devilee P, Meindl A, Couch FJ, Southey M, Goldgar DE, Evans DG, Chenevix-Trench G, Rahman N, Robson M, Domchek SM, Foulkes WD. Gene-panel sequencing and the prediction of breast-cancer risk. *N Engl J Med* 2015;372:2243–57.

17 Caudle KE, Dunnenberger HM, Freimuth RR, Peterson JF, Burlison JD, Whirl-Carrillo M, Scott SA, Rehm HL, Williams MS, Klein TE, Relling MV, Hoffman JM. Standardizing terms for clinical pharmacogenetic test results: consensus terms from the Clinical Pharmacogenetics Implementation Consortium (CPIC). *Genet Med* 2017;19:215–23.

18 Quax TE, Claassens NJ, Söll D, van der Oost J. Codon Bias as a Means to Fine-Tune Gene Expression. *Mol Cell* 2015;59:149–61.

19 Plotkin JB, Kudla G. Synonymous but not the same: the causes and consequences of codon bias. *Nat Rev Genet* 2011;12:32–42.

20 Miller JE, Shivakumar MK, Risacher SL, Saykin AJ, Lee S, Nho K, Kim D. Codon bias among synonymous rare variants is associated with Alzheimer's disease imaging biomarker. *Pac Symp Biocomput* 2018;23:365–76.

21 Pollard KS, Hubisz MJ, Rosenbloom KR, Siepel A. Detection of nonneutral substitution rates on mammalian phylogenies. *Genome Res* 2010;20:110–21.

22 Nalpathamkalam T, Derkach A, Paterson AD, Merico D. Genetic Analysis Workshop 18 single-nucleotide variant prioritization based on protein impact, sequence conservation, and gene annotation. *BMC Proc* 2014;8(Suppl 1 Genetic Analysis Workshop 18Vanessa Olmo):S11.

23 Stormo GD, Schneider TD, Gold L, Ehrenfeucht A. Use of the 'Perceptron' algorithm to distinguish translational initiation sites in E. coli. *Nucleic Acids Res* 1982;10:2997–3011.

24 Yeo G, Burge CB. Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals. *J Comput Biol* 2004;11(2-3):377–94.

25 Tavtigian SV, Greenblatt MS, Lesueur F, Byrnes GB. IARC Unclassified Genetic Variants Working Group. In silico analysis of missense substitutions using sequence-alignment based methods. *Hum Mutat* 2008;29:1327–36.

26 Mathe E, Olivier M, Kato S, Ishioka C, Hainaut P, Tavtigian SV. Computational approaches for predicting the biological effect of p53 missense mutations: a comparison of three sequence analysis based methods. *Nucleic Acids Res* 2006;34:1317–25.

27 Tavtigian SV, Byrnes GB, Goldgar DE, Thomas A. Classification of rare missense substitutions, using risk surfaces, with genetic- and molecular-epidemiology applications. *Hum Mutat* 2008;29:1342–54.

28 Vallée MP, Di Sera TL, Nix DA, Paquette AM, Parsons MT, Bell R, Hoffman A, Hogervorst FB, Goldgar DE, Spurdle AB, Tavtigian SV. Adding In Silico Assessment of Potential Splice Aberration to the Integrated Evaluation of BRCA Gene Unclassified Variants. *Hum Mutat* 2016;37:627–39.

29 Thompson BA, Greenblatt MS, Vallee MP, Herkert JC, Tessereau C, Young EL, Adzhubey IA, Li B, Bell R, Feng B, Mooney SD, Radivojac P, Sunyaev SR, Frebourg T, Hofstra RM, Sijmons RH, Boucher K, Thomas A, Goldgar DE, Spurdle AB, Tavtigian SV. Calibration of multiple in silico tools for predicting pathogenicity of mismatch repair gene missense substitutions. *Hum Mutat* 2013;34:255–65.

30 Fortuno C, James PA, Young EL, Feng B, Olivier M, Pesaran T, Tavtigian SV, Spurdle AB. Improved, ACMG-compliant, in silico prediction of pathogenicity for missense substitutions encoded by TP53 variants. *Hum Mutat* 2018;39:1061–9.

31 Kim MP, Zhang Y, Lozano G. Mutant p53: Multiple Mechanisms Define Biologic Activity in Cancer. *Front Oncol* 2015;5:249.

32 Raue F, Frank-Raue K. Update multiple endocrine neoplasia type 2. *Fam Cancer* 2010;9:449–57.

33 de la Hoya M, Soukarieh O, López-Perolio I, Vega A, Walker LC, van Ierland Y, Baralle D, Santamariña M, Lattimore V, Wijnen J, Whiley P, Blanco A, Raponi M, Hauke J, Wappenschmidt B, Becker A, Hansen TV, Behar R, Investigators K, Niederacher D, Arnold N, Dworniczak B, Steinemann D, Faust U, Rubinstein W, Hulick PJ, Houdayer C, Caputo SM, Castera L, Pesaran T, Chao E, Brewer C, Southey MC, van Asperen CJ, Singer CF, Sullivan J, Poplawski N, Mai P, Peto J, Johnson N, Burwinkel B, Surowy H, Bojesen SE, Flyger H, Lindblom A, Margolin S, Chang-Claude J, Rudolph A, Radice P, Galastri L, Olson JE, Hallberg E, Giles GG, Milne RL, Andrulis IL, Glendon G, Hall P, Czene K, Blows F, Shah M, Wang Q, Dennis J, Michailidou K, McGuffog L, Bolla MK, Antoniou AC, Easton DF, Couch FJ, Tavtigian S, Vreeswijk MP, Parsons M, Meeks HD, Martins A, Goldgar DE, Spurdle AB. Combined genetic and splicing analysis of BRCA1 c.[594-2A>C; 641A>G] highlights the relevance of naturally occurring in-frame transcripts for developing disease gene variant classification algorithms. *Hum Mol Genet* 2016;25:2256–68.

34 Colombo M, Lòpez-Perolio I, Meeks HD, Caleca L, Parsons MT, Li H, De Vecchi G, Tudini E, Foglia C, Mondini P, Manoukian S, Behar R, Garcia EBG, Meindl A, Montagna M, Niederacher D, Schmidt AY, Varesco L, Wappenschmidt B, Bolla MK, Dennis J, Michailidou K, Wang Q, Aittomäki K, Andrulis IL, Anton-Culver H, Arndt V, Beckmann MW, Beeghly-Fadel A, Benitez J, Boeckx B, Bogdanova NV, Bojesen SE, Bonanni B, Brauch H, Brenner H, Burwinkel B, Chang-Claude J, Conroy DM, Couch FJ, Cox A, Cross SS, Czene K, Devilee P, Dörk T, Eriksson M, Fasching PA, Figueroa J, Fletcher O, Flyger H, Gabrielson M, García-Closas M, Giles GG, González-Neira A, Guénel P, Haiman CA, Hall P, Hamann U, Hartman M, Hauke J, Hollestelle A, Hopper JL, Jakubowska A, Jung A, Kosma VM, Lambrechts D, Le Marchand L, Lindblom A, Lubinski J, Mannermaa A, Margolin S, Miao H, Milne RL, Neuhausen SL, Nevanlinna H, Olson JE, Peterlongo P, Peto J, Pylkäs K, Sawyer EJ, Schmidt MK, Schmutzler RK, Schneeweiss A, Schoemaker MJ, See MH, Southey MC, Swerdlow A, Teo SH, Toland AE, Tomlinson I, Truong T, van Asperen CJ, van den Ouweland AMW, van der Kolk LE, Winqvist R, Yannoukakos D, Zheng W, Dunning AM, Easton DF, Henderson A, Hogervorst FBL, Izatt L, Offitt K, Side LE, van Rensburg EJ, Embrace S, Hebon S, McGuffog L, Antoniou AC, Chenevix-Trench G, Spurdle AB, Goldgar DE, Hoya M, Radice P. kConFab/AOCS Investigators. The BRCA2 c.68-7T > A variant is not pathogenic: A model for clinical calibration of spliceogenicity. *Hum Mutat* 2018;39:729–41.

35 Jager KJ, Zoccali C, Kramar R, Dekker FW, occurrence Mdisease *Kidney Int* 2007;72:412–5.

36 Thompson BA, Spurdle AB, Plazzer JP, Greenblatt MS, Akagi K, Al-Mulla F, Bapat B, Bernstein I, Capellá G, den Dunnen JT, du Sart D, Fabre A, Farrell MP, Farrington SM,

Frayling IM, Frebourg T, Goldgar DE, Heinen CD, Holinski-Feder E, Kohonen-Corish M, Robinson KL, Leung SY, Martins A, Moller P, Morak M, Nystrom M, Peltomaki P, Pineda M, Qi M, Ramesar R, Rasmussen LJ, Royer-Pokora B, Scott RJ, Sijmons R, Tavtigian SV, Tops CM, Weber T, Wijnen J, Woods MO, Macrae F, Genuardi M. Application of a 5-tiered scheme for standardized classification of 2,360 unique mismatch repair gene variants in the InSiGHT locus-specific database. *Nat Genet* 2014;46:107–15.

37 Meeks HD, Song H, Michailidou K, Bolla MK, Dennis J, Wang Q, Barrowdale D, Frost D, McGuffog L, Ellis S, Feng B, Buys SS, Hopper JL, Southey MC, Tesoriero A, James PA, Bruinsma F, Campbell IG, Broeks A, Schmidt MK, Hogervorst FB, Beckman MW, Fasching PA, Fletcher O, Johnson N, Sawyer EJ, Riboli E, Banerjee S, Menon U, Tomlinson I, Burwinkel B, Hamann U, Marme F, Rudolph A, Janavicius R, Tihomirova L, Tung N, Garber J, Cramer D, Terry KL, Poole EM, Tworoger SS, Dorfling CM, van Rensburg EJ, Godwin AK, Guénel P, Truong T, Stoppa-Lyonnet D, Damiola F, Mazoyer S, Sinilnikova OM, Isaacs C, Maugard C, Bojesen SE, Flyger H, Gerdes AM, Hansen TV, Jensen A, Kjaer SK, Hogdall C, Hogdall E, Pedersen IS, Thomassen M, Benitez J, González-Neira A, Osorio A, Hoya ML, Segura PP, Diez O, Lazaro C, Brunet J, Anton-Culver H, Eunjung L, John EM, Neuhausen SL, Ding YC, Castillo D, Weitzel JN, Ganz PA, Nussbaum RL, Chan SB, Karlan BY, Lester J, Wu A, Gayther S, Ramus SJ, Sieh W, Whittemore AS, Monteiro AN, Phelan CM, Terry MB, Piedmonte M, Offit K, Robson M, Levine D, Moysich KB, Cannioto R, Olson SH, Daly MB, Nathanson KL, Domchek SM, Lu KH, Liang D, Hildebrant MA, Ness R, Modugno F, Pearce L, Goodman MT, Thompson PJ, Brenner H, Butterbach K, Meindl A, Hahnen E, Wappenschmidt B, Brauch H, Brüning T, Blomqvist C, Khan S, Nevanlinna H, Pelttari LM, Aittomäki K, Butzow R, Bogdanova NV, Dörk T, Lindblom A, Margolin S, Rantala J, Kosma VM, Mannermaa A, Lambrechts D, Neven P, Claes KB, Maerken TV, Chang-Claude J, Flesch-Janys D, Heitz F, Varon-Mateeva R, Peterlongo P, Radice P, Viel A, Barile M, Peissel B, Manoukian S, Montagna M, Oliani C, Peixoto A, Teixeira MR, Collavoli A, Hallberg E, Olson JE, Goode EL, Hart SN, Shimelis H, Cunningham JM, Giles GG, Milne RL, Healey S, Tucker K, Haiman CA, Henderson BE, Goldberg MS, Tischkowitz M, Simard J, Soucy P, Eccles DM, Le N, Borresen-Dale AL, Kristensen V, Salvesen HB, Bjorge L, Bandera EV, Risch H, Zheng W, Beeghly-Fadiel A, Cai H, Pylkäs K, Tollenaar RA, Ouweland AM, Andrulis IL, Knight JA, Narod S, Devilee P, Winqvist R, Figueroa J, Greene MH, Mai PL, Loud JT, García-Closas M, Schoemaker MJ, Czene K, Darabi H, McNeish I, Siddiquil N, Glasspool R, Kwong A, Park SK, Teo SH, Yoon SY, Matsuo K, Hosono S, Woo YL, Gao YT, Foretova L, Singer CF, Rappaport-Feurhauser C, Friedman E, Laitman Y, Rennert G, Imyanitov EN, Hulick PJ, Olopade OI, Senter L, Olah E, Doherty JA, Schildkraut J, Koppert LB, Kiemeney LA, Massuger LF, Cook LS, Pejovic T, Li J, Borg A, Öfverholm A, Rossing MA, Wentzensen N, Henriksson K, Cox A, Cross SS, Pasini BJ, Shah M, Kabisch M, Torres D, Jakubowska A, Lubinski J, Gronwald J, Agnarsson BA, Kupryjanczyk J, Moes-Sosnowska J, Fostira F, Konstantopoulou I, Slager S, Jones M, Antoniou AC, Berchuck A, Swerdlow A, Chenevix-Trench G, Dunning AM, Pharoah PD, Hall P, Easton DF, Couch FJ, Spurdle AB, Goldgar DE. EMBRACE kConFab Investigators Australia Ovarian Cancer Study Group HEBON GEMO Study Collaborators OCGN PRostate cancer AssoCiation group To Investigate Cancer Associated aLterations in the genome. BRCA2 Polymorphic Stop Codon K3326X and the Risk of Breast, Prostate, and Ovarian Cancers. *J Natl Cancer Inst* 2016;108:djv315.

38 Arbustini E, Sgarella A, Ferrari A, Grasso D, Cassani C, Lucioni M, Di Giulio G, Grasso M. RE: BRCA2 Polymorphic Stop Codon K3326X and the Risk of Breast, Prostate, and Ovarian Cancers. *J Natl Cancer Inst* 2016;108:djw172.

39 Spurdle AB, Domchek S, Robson M, Buys S, Radice P, de la Hoya M, Devilee P, Monteiro AN, Southey M, Eccles D, Couch FJ, Goldgar DE. all authors. Response. *J Natl Cancer Inst* 2016;108.

40 Patel RY, Shah N, Jackson AR, Ghosh R, Pawliczek P, Paithankar S, Baker A, Riehle K, Chen H, Milosavljevic S, Bizon C, Rynearson S, Nelson T, Jarvik GP, Rehm HL, Harrison SM, Azzariti D, Powell B, Babb L, Plon SE, Milosavljevic A. ClinGen Resource. ClinGen Pathogenicity Calculator: a configurable system for assessing pathogenicity of genetic variants. *Genome Med* 2017;9:3.

41 Ioannidis NM, Rothstein JH, Pejaver V, Middha S, McDonnell SK, Baheti S, Musolf A, Li Q, Holzinger E, Karyadi D, Cannon-Albright LA, Teerlink CC, Stanford JL, Isaacs WB, Xu J, Cooney KA, Lange EM, Schleutker J, Carpten JD, Powell IJ, Cussenot O, Cancel-Tassin G, Giles GG, MacInnis RJ, Maier C, Hsieh CL, Wiklund F, Catalona WJ, Foulkes WD, Mandal D, Eeles RA, Kote-Jarai Z, Bustamante CD, Schaid DJ, Hastie T, Ostrander EA, Bailey-Wilson JE, Radivojac P, Thibodeau SN, Whittemore AS, Sieh W. REVEL: An Ensemble Method for Predicting the Pathogenicity of Rare Missense Variants. *Am J Hum Genet* 2016;99:877–85.

42 Kircher M, Witten DM, Jain P, O'Roak BJ, Cooper GM, Shendure J. A general framework for estimating the relative pathogenicity of human genetic variants. *Nat Genet* 2014;46:310–5.

43 Feng BJ. PERCH: A Unified Framework for Disease Gene Prioritization. *Hum Mutat* 2017;38:243–51.

44 Lee A, Mavaddat N, Wilcox AN, Cunningham AP, Carver T, Hartley S, Babb de Villiers C, Izquierdo A, Simard J, Schmidt MK, Walter FM, Chatterjee N, Garcia-Closas M, Tischkowitz M, Pharoah P, Easton DF, Antoniou AC. BOADICEA: a comprehensive breast cancer risk prediction model incorporating genetic and nongenetic risk factors. *Genet Med* 2019 [Epub ahead of print 15 Jan 2019].