Journal of
Medical Genetics

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

Journal:	*Journal of Medical Genetics*
Manuscript ID	jmedgenet-2018-105872
Article Type:	Position Statement
Date Submitted by the Author:	13-Nov-2018
Complete List of Authors:	Spurdle, Amanda; Queensland Institute of Medical Research, Genetics and Population Health Division Greville-Haygate, Stephanie; University of Southampton, Faculty of Medicine Antoniou, Antonis; University of Cambridge, Centre for Cancer Genetic Epidemiology Brown, Melissa; University of Queensland, School of Chemistry and Molecular Biosciences Burke, Leslie; University of Queensland, School of Chemistry and Molecular Biosciences de la Hoya, Miguel; Hospital Clínico San Carlos, Medical Oncology Department Domchek, Susan; Basser Center for BRCA, Abramson Cancer Center, Perelman School of Medicine at the University of Pennsylvania, Dork, Thilo; Hannover Medical School, Gynaecology Research Unit Firth, Helen; Cambridge University Hospitals National Health Service Foundation Trust Monteiro, Alvaro; H. Lee Moffitt Cancer Center and Research Institute, Cancer Epidemiology Mensenkamp, Arjen; Radboud University Medical Center, Department of Human Genetics Parsons, Michael; Queensland Institute of Medical Research, Genetics and Population Health Division Radice, Paolo; Fondazione IRCCS Istituto Nazionale dei Tumori, Preventive and Predictive Medicine Robson, Mark; Memorial Sloan Kettering Cancer Center, Clinical Genetics Service Tischkowitz, Marc; Cambridge University , Department of Medical Genetics Tudini, Emma; Queensland Institute of Medical Research, Genetics and Population Health Division turnbull, clare; Institute of Cancer Research, Cancer Genetics Vreeswijk, Maaike; Leiden University Medical Center Walker, Logan; University of Otago, Department of Pathology; University of Otago, Christchurch, Tavtigian, Sean; Huntsman Cancer Institute, Department of Oncological Sciences Eccles, Diana; University of Southampton, Faculty of Medicine

Keywords:	genetic variant, interpretation, reporting, vocabulary, cancer susceptibility

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 A. Brown[4], Leslie J. Burke[4], Miguel de la Hoya[5], Susan M. Domchek[6], Thilo Dörk[7], Helen Firth[8,9], Alvaro N. Monteiro[10], Arjen R. Mensenkamp[11], Michael T. Parsons[1] Paolo Radice[12], Mark Robson[13], Marc Tischkowitz[14], Emma Tudini[1], Clare Turnbull[15,16], Maaike Vreeswijk[17], Logan Walker[18], Sean Tavtigian[19,20], Diana Eccles[2]*

1. QIMR Berghofer Medical Research Institute, Herston, Brisbane, 4006, Australia
2. Faculty of Medicine University of Southampton, Tremona Road, Southampton SO16 6YD, UK.
3. Centre for Cancer Genetic Epidemiology, Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK.
4. School of Chemistry and Molecular Biosciences, University of Queensland, Brisbane, 4072, Australia.
5. Medical Oncology Laboratory, Hospital Clínico San Carlos. Instituto de Investigación Sanitaria San Carlos (IdISSC), Centro Investigación Biomédica en Red de Cáncer (CIBERONC), Madrid, 28040, Spain.
6. Basser Center for BRCA, Abramson Cancer Center, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, PA, USA.
7. Hannover Medical School, Gynaecology Research Unit, 30625, Hannover, Germany.
8. Cambridge University Hospitals National Health Service Foundation Trust, Cambridge Biomedical Campus, Hills Road, Cambridge CB2 0QQ, UK.
9. Wellcome Trust Sanger Institute, Wellcome Genome Campus, Cambridge, Hinxton CB10 1SA, UK.
10. Cancer Epidemiology Program, H. Lee Moffitt Cancer Center, Tampa, FL, USA.
11. Radboud University Medical Centre, Nijmegen, the Netherlands.
12. Unit of Molecular Bases of Genetic Risk and Genetic Testing, Department of Research, Fondazione IRCCS (Istituto Di Ricovero e Cura a Carattere Scientifico) Istituto Nazionale dei Tumori (INT), Milan, Italy.
13. Clinical Genetics Service, Department of Medicine, Memorial Sloan-Kettering Cancer Center, New York, NY, USA.
14. Department of Medical Genetics, Cambridge University, Cambridge, UK.
15. Division of Genetics and Epidemiology, Institute of Cancer Research, London, UK.
16. William Harvey Research Institute, Queen Mary University, London, UK.
17. Department of Human Genetics, Leiden University Medical Center, PO Box 9600, Leiden, 2300RC, The Netherlands.
18. Department of Pathology and Biomedical Science, University of Otago, Christchurch, New Zealand.
19. Department of Oncological Sciences, University of Utah School of Medicine, Salt Lake City, United States.
20. Huntsman Cancer Institute, University of Utah, Salt Lake City, United States.

* Co-corresponding authors.
Amanda B. Spurdle, PhD, Group Leader, QIMR Berghofer Medical Research Institute, 300 Herston Rd, Herston, Brisbane, 4006, Australia; Amanda.Spurdle@qimrberghofer.edu.au; +61 (0)7 3362 0371
Diana Eccles, postal address, e-mail and telephone number.
Professor Diana Eccles MD FRCP, Dean of Medicine, Faculty of Medicine, University Of Southampton, Room AB215, MP801, Level B, South Academic Block, University Hospital Southampton,Tremona Road, Southampton, SO16 6YD; D.M.Eccles@soton.ac.uk; +44 (0)23 8120 6581

Word count, excluding title page, abstract, references, figures and tables: 4995

Keywords: genetic variant, interpretation, reporting, vocabulary, cancer susceptibility

SUMMARY

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 tumor 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 amongst 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 ENIGMA consortium (Evidence-Based Network for the Interpretation of Germline Mutant Alleles) 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 **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 multi-cancer 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- characterized 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 multi-gene 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 to the "average pathogenic" variant for that gene. Different terms in use were considered by ENIGMA members attending the June 2016 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 multi-tier 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 standardized terminology and definitions for describing sequence variation, focused on inherited variants

Supplementary Table 1 summarizes 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 i.e. 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 tumorigenesis. Indeed, the interface of the CIViC (Clinical Interpretation of Variants in Cancer) knowledgebase [9] describes variants for a specific gene using the term "mutation", with additional qualifications e.g. 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)".

Further, 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 recognized.

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 (EMA; http://www.ema.europa.eu/) and the US Food & Drug administration (FDA, https://www.fda.gov/) to denote eligibility of patients with specified cancer types/presentation for PARP inhibitor therapy, namely patients with a "deleterious or suspected deleterious germline [or somatic] BRCA mutation". Further, 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 familial breast-ovarian cancer patients. 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 summarizing all the component information used for classification, is thus paramount to avoid variant misinterpretation and inappropriate patient management.

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

Program	Output terms and other descriptions
CONDEL (CONsensus DELeteriousness score of missense mutations)	**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.
FATHMM (Functional Analysis through Hidden Markov Models)	Damaging Tolerated http://fathmm.biocompute.org.uk
LRT (Likelihood ratio test)	**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
nsSNPAnalyzer (predicting disease associated non-synonymous single nucleotide polymorphisms)	Disease Neutral http://snpanalyzer.uthsc.edu/
PhD-SNP (Predictor of human Deleterious Single Nucleotide Polymorphisms)	Disease Neutral http://snps.biofold.org/phd-snp/
PolyPhen-2 (Polymorphism Phenotyping v2)	Probably Damaging Possibly Damaging **Benign** http://genetics.bwh.harvard.edu/pph2
Provean (Protein Variation Effect Analyzer)	**Deleterious** Neutral http://provean.jcvi.org
SIFT (Sorting Intolerant From Tolerant)	Damaging Tolerated http://sift-dna.org

* NOTE: Prediction tools used for missense variants highlighted in grey are included as options for scoring bioinformatic predictions in the ClinGen ACMG calculator [13]. The preferred ClinGen meta-predictor REVEL (Rare Exome Variant Ensemble Learner) [14], and other meta-predictors CADD (Combined Annotation Dependent Depletion) [15], and BayesDel [16], provide continuous scores and not specific terms as output. Output terms also used in clinical reporting, or to define eligibility for PARP inhibitor treatment, are noted in **bold**.

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 tumors in the context of drug therapy selection for cancer patients, or distinguishing true germline variants from variants arising from somatic clonal drift in "disease free" tissue used for DNA extraction [17 18]. These suggestions take into consideration terms put forward by the IARC unclassified sequence variants working group [6], ACMG/AMP [5] and HVP [19], and a comprehensive review article assessing clinical implications of gene panel test results for breast cancer risk prediction [20]. 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 [21], 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 [19]. 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 (e.g. 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 tumor DNA (tumor biopsy or circulating tumor DNA derived from blood), the variant will be present in tumor 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 tumor sequencing (**tumor-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 hematopoetic cells (particularly *TP53* [17 18]), or from circulating tumor 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 [22-24], 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 analyzed by evolutionary/phylogenetic methods involve alignment of at least three nucleic acid sequences, termed multiple (multi-species) sequence alignment (acronym MSA). We suggest that such analysis specify the method/program 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:

- **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 [25], may be helpful to assess if a specific position is evolutionary constrained or not [26].

Further, position weight matrices [27] developed for functionally important sequence motifs eg splice junctions [28] 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 [29].

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
- **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 a multiple sequence alignment. 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)[30].

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.
- **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 e.g. 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 e.g. 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 [31-34]. See **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 analyzed 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 i.e. any combination of mRNA transcripts predicted non-coding, predicted protein truncating-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 i.e. 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* i.e. 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 analyzed 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 to 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) i.e. 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 e.g for p53 [35], RET [36]. 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 variant protein that disrupts protein-protein dimerization or oligomerization or interaction with a co-factor or interacting protein

Note: A variant with measurable effect *in vitro* on mRNA transcript profile or protein function (specifying feature measured), relative to appropriate controls, should not 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 [37] for example calibration of *BRCA1* 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 [38]. 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 e.g. 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 e.g. 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 to 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 (e.g. 50 in every 100 people with this variant (type) are expected to develop breast or ovarian cancer by age 70) and a relative measure (e.g. a variant carrier is 10 times more likely to develop breast cancer in their lifetime compared to women in the general population), and report these with appropriate confidence intervals. Based on descriptors applied previously for breast cancer [20], for this discussion document we have categorized cancer risk levels associated with a given variant, relative to the general population risk, as follows: **High increased risk,** >4-fold; **Moderate increased risk,** 2-4-fold; **Low increased risk,** <2-fold. 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 5-tier system

The IARC 5-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 [39], 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 (See **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 less than 1.5-fold increased risk of breast or ovarian cancer [40], both publicly [41 42], and by direct query to the BRCA-Exchange website (http://brcaexchange.org/). The latter led to a change in representation of this tier as "benign" to "benign-little clinical significance" on the BRCA-Exchange website.

Further, during development of the ENIGMA *BRCA1/2* variant classification criteria (https://enigmaconsortium.org), research results emphasized 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 5-tier system, in particular if the disease penetrance was sufficient to trigger altered management albeit not as extensive as the "standard pathogenic" variant for that gene [3]. The advent of multi-gene 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.

Table 2: Alternative terms currently in use to describe 5-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**
Class 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 health-care provider can use the molecular testing information in clinical decision making
Class 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 health-care provider can use the molecular testing information in clinical decision making when combined with other evidence of the disease in question.
Class 3	Uncertain†	0.050-0.949	No (recommend **research** testing of family members)	Based on family history & 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
Class 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 (i.e.based on family history & other risk factors)	Likely Benign	0.001-0.099	Sufficient evidence that a health-care provider can conclude that it is not the cause of the patient's disorder when combined with other information
Class 1	Not pathogenic or of **no** clinical significance§	<0.001	No	Treat as "no pathogenic variant detected" for this disorder (i.e.based on family history & other risk factors)	Benign	<0.001	Sufficient evidence that a health-care 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
** 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/AMP guidelines do not require quantitative variant classification methods to be but nevertheless propose probabilities of a variant either being disease-causing or benign [5].

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 **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 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 5 tiers intended to capture the complexity of reporting in the multi-gene 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 5-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.

Table 3: Recommended terminology and descriptors for 5-tier disease gene variant classification categories, considering variant pathogenicity in the multi-gene panel testing arena.*

Numeric Class	Consolidated 5-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 (>4-fold risk relative to population)	***Moderate penetrant variants (2-4-fold relative risk)***
Class 5	Pathogenic	P	Sequence variant is associated with ≥2-fold 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
Class 4	Likely Pathogenic	LP	Sequence variant is likely** associated with ≥2-fold 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.***
Class 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.	
Class 2	Likely Little clinical significance/ Likely Benign†	LB	Sequence variant is likely NOT associated with ≥2-fold 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.	
Class 1	Little clinical significance/ Benign†	B	Sequence Variant is NOT associated with ≥2-fold 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 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 >2-fold 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 <2-fold 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 [43]. Further implementation research is required to understand how best to implement PRS testing to stratify cancer risks in a range of settings, including cancer patients 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 e.g. the *BRCA2* c.9976A>T p.(Lys3326Ter) variant associated with less than 1.5-fold increased risk of breast or ovarian cancer [40].

Proposal for development of a multi-tier system for variant annotation in clinical test reporting of multi-gene panel results

Despite the expansion of descriptions for the 5-tier variant classification system shown in Table 3, it was clear from comments received that assignment of variant pathogenicity using the current 5 tier system is inadequate to deal with the complexities of reporting multi-gene panel testing outcomes, and to portray differences in variant-specific risks for a given gene. The term “pathogenic” remained contentious, with comments raised including: need to capture the relevance of genetic findings to patient disease diagnosis (phenotype) versus relevance of a secondary finding (i.e. 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 minimize ambiguity of clinical management considerations. We thus developed a template to emphasize the value of a multi-tier reporting system, (outlined in **Table 4**), and provide several worked examples (**Supplementary Table 2**) to indicate its potential to capture the complexity of clinical actionability for variants identified by multi-gene 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 multi-gene panel reporting.

Table 4: Suggested approach to multi-tier 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
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 endpoints 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 Details of the evidence supporting the variant classification is 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 (>4 fold) 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 evidence to support a moderate (2-4 fold) 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 evidence from population studies to support that is variant is associated with a low (<2-fold) 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 & Actionability*	This final element comprises the discussion between physician and patient. It may take the form of a personalised assessment of risk based upon the presenting cancer phenotype, clinical scenario and family history. 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 Upon a Comprehensive Risk Evaluation + Additional moderate 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 Upon a Comprehensive Risk Evaluation + : Additional moderate 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 Upon 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.
+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.	

* 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 cancer patient 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 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 multi-tier reporting schemes provide for distinct annotation of germline variants for disease risk and relevance to drug treatment.

CONCLUSIONS & 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. Further, 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 summarize the key points and provide recommendations on variant annotation and terminology in **Table 5**. 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 recognize that practical implementation of such a system would require that there is 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.

Table 5: Key recommendations regarding variant descriptors and their use in 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 analyzed in **functional assays** (biochemical, biophysical, molecular biological, cellular) that assess variant effect on protein conformation/stability/activity/function should describe effect compared to wildtype 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 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 confidence intervals.

* Note, **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. Please see **Supplementary Text** for an overview of variant annotation.

ACKNOWLEDGEMENTS

We thank ENIGMA collaborators for helpful feedback provided at ENIGMA general meetings, and other verbal discussion.

COMPETING INTERESTS

Author M Robson discloses the following support: Honoraria (Advisory) from AstraZeneca, Pfizer; Consulting or Advisory from McKesson, AstraZeneca; Research Funding from AstraZeneca (Institution), Myriad (Institution, in-kind), Invitae (Institution, in-kind), Pfizer (institution), AbbVie (institution), Tesaro (institution), Medivation (Institution); Travel, Accommodation, Expenses from AstraZeneca.

Susan Domchek discloses the following support: Honoraria (Advisory) from AstraZeneca, Clovis, and Bristol Myers Squibb.

FUNDING

ABS is supported by an Australian National Health and Medical Research Council (NHMRC) Senior Research Fellowship (ID1061779). ACA is supported by Cancer Research UK (C12292/A20861). MAB and LJB have been supported by the Australian NHMRC (ID1104808) and the Cancer Council Queensland (ID1044008, ID1026095). SMD is supported by the Breast Cancer Research Foundation and Komen. MdH is supported by funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 634935 (BRIDGES), and by Spanish Instituto de Salud Carlos III (ISCIII) funding (grants PI15/00059), an initiative of the Spanish Ministry of Economy and Innovation partially supported by European Regional Development FEDER Funds. MP was supported in part by NHMRC (ID1101400). MR was funded in part through an NIH/NCI Cancer Center Support Grant P30 CA008748. ET was supported by the Australian NHMRC (ID1104808). LCW is supported by a Rutherford Discovery Fellowship (Royal Society of New Zealand).

References:

1. Spurdle AB, Couch FJ, Parsons MT, et al. 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**(6):3419 doi: 10.1186/s13058-014-0474-y[published Online First: Epub Date]|.
2. Spurdle AB, Whiley PJ, Thompson B, et al. BRCA1 R1699Q variant displaying ambiguous functional abrogation confers intermediate breast and ovarian cancer risk. J Med Genet 2012;**49**(8):525-32 doi: 10.1136/jmedgenet-2012-101037[published Online First: Epub Date]|.
3. Moghadasi S, Meeks HD, Vreeswijk MP, et al. 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**(1):15-20 doi: 10.1136/jmedgenet-2017-104560[published Online First: Epub Date]|.
4. Eccles DM, Mitchell G, Monteiro AN, et al. BRCA1 and BRCA2 genetic testing-pitfalls and recommendations for managing variants of uncertain clinical significance. Ann Oncol 2015;**26**(10):2057-65 doi: 10.1093/annonc/mdv278[published Online First: Epub Date]|.
5. Richards S, Aziz N, Bale S, et al. 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**(5):405-24 doi: 10.1038/gim.2015.30[published Online First: Epub Date]|.
6. Plon SE, Eccles DM, Easton D, et al. Sequence variant classification and reporting: recommendations for improving the interpretation of cancer susceptibility genetic test results. Hum Mutat 2008;**29**(11):1282-91 doi: 10.1002/humu.20880[published Online First: Epub Date]|.
7. MacArthur DG, Manolio TA, Dimmock DP, et al. Guidelines for investigating causality of sequence variants in human disease. Nature 2014;**508**(7497):469-76 doi: 10.1038/nature13127[published Online First: Epub Date]|.
8. Chenevix-Trench G, Milne RL, Antoniou AC, et al. 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**(2):104 doi: 10.1186/bcr1670[published Online First: Epub Date]|.
9. Griffith M, Spies NC, Krysiak K, et al. CIViC is a community knowledgebase for expert crowdsourcing the clinical interpretation of variants in cancer. Nat Genet 2017;**49**(2):170-74 doi: 10.1038/ng.3774[published Online First: Epub Date]|.
10. Fokkema IF, Taschner PE, Schaafsma GC, et al. LOVD v.2.0: the next generation in gene variant databases. Hum Mutat 2011;**32**(5):557-63 doi: 10.1002/humu.21438[published Online First: Epub Date]|.
11. Lindor NM, Guidugli L, Wang X, et al. A review of a multifactorial probability-based model for classification of BRCA1 and BRCA2 variants of uncertain significance (VUS). Hum Mutat 2012;**33**(1):8-21 doi: 10.1002/humu.21627[published Online First: Epub Date]|.

12. Adzhubei IA, Schmidt S, Peshkin L, et al. A method and server for predicting damaging missense mutations. Nat Methods 2010;**7**(4):248-9 doi: 10.1038/nmeth0410-248[published Online First: Epub Date]|.
13. Patel RY, Shah N, Jackson AR, et al. ClinGen Pathogenicity Calculator: a configurable system for assessing pathogenicity of genetic variants. Genome Med 2017;**9**(1):3 doi: 10.1186/s13073-016-0391-z[published Online First: Epub Date]|.
14. Ioannidis NM, Rothstein JH, Pejaver V, et al. REVEL: An Ensemble Method for Predicting the Pathogenicity of Rare Missense Variants. Am J Hum Genet 2016;**99**(4):877-85 doi: 10.1016/j.ajhg.2016.08.016[published Online First: Epub Date]|.
15. Kircher M, Witten DM, Jain P, et al. A general framework for estimating the relative pathogenicity of human genetic variants. Nat Genet 2014;**46**(3):310-5 doi: 10.1038/ng.2892[published Online First: Epub Date]|.
16. Feng BJ. PERCH: A Unified Framework for Disease Gene Prioritization. Hum Mutat 2017;**38**(3):243-51 doi: 10.1002/humu.23158[published Online First: Epub Date]|.
17. Coffee B, Cox HC, Kidd J, et al. Detection of somatic variants in peripheral blood lymphocytes using a next generation sequencing multigene pan cancer panel. Cancer Genet 2017;**211**:5-8 doi: 10.1016/j.cancergen.2017.01.002[published Online First: Epub Date]|.
18. Weitzel JN, Chao EC, Nehoray B, et al. Somatic TP53 variants frequently confound germ-line testing results. Genet Med 2017 doi: 10.1038/gim.2017.196[published Online First: Epub Date]|.
19. Project) HHV. HVP Guideline HVP/GL/003-01/EN, 2016.
20. Easton DF, Pharoah PD, Antoniou AC, et al. Gene-panel sequencing and the prediction of breast-cancer risk. N Engl J Med 2015;**372**(23):2243-57 doi: 10.1056/NEJMsr1501341[published Online First: Epub Date]|.
21. Caudle KE, Dunnenberger HM, Freimuth RR, et al. Standardizing terms for clinical pharmacogenetic test results: consensus terms from the Clinical Pharmacogenetics Implementation Consortium (CPIC). Genet Med 2017;**19**(2):215-23 doi: 10.1038/gim.2016.87[published Online First: Epub Date]|.
22. Quax TE, Claassens NJ, Soll D, et al. Codon Bias as a Means to Fine-Tune Gene Expression. Mol Cell 2015;**59**(2):149-61 doi: 10.1016/j.molcel.2015.05.035[published Online First: Epub Date]|.
23. Plotkin JB, Kudla G. Synonymous but not the same: the causes and consequences of codon bias. Nat Rev Genet 2011;**12**(1):32-42 doi: 10.1038/nrg2899[published Online First: Epub Date]|.
24. Miller JE, Shivakumar MK, Risacher SL, et al. Codon bias among synonymous rare variants is associated with Alzheimer's disease imaging biomarker. Pac Symp Biocomput 2018;**23**:365-76
25. Pollard KS, Hubisz MJ, Rosenbloom KR, et al. Detection of nonneutral substitution rates on mammalian phylogenies. Genome Res 2010;**20**(1):110-21 doi: 10.1101/gr.097857.109[published Online First: Epub Date]|.
26. Nalpathamkalam T, Derkach A, Paterson AD, et al. 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 doi: 10.1186/1753-6561-8-S1-S11[published Online First: Epub Date]|.

27. Stormo GD, Schneider TD, Gold L, et al. Use of the 'Perceptron' algorithm to distinguish translational initiation sites in E. coli. Nucleic Acids Res 1982;**10**(9):2997-3011
28. 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 doi: 10.1089/1066527041410418[published Online First: Epub Date]|.
29. Tavtigian SV, Greenblatt MS, Lesueur F, et al. In silico analysis of missense substitutions using sequence-alignment based methods. Hum Mutat 2008;**29**(11):1327-36 doi: 10.1002/humu.20892[published Online First: Epub Date]|.
30. Mathe E, Olivier M, Kato S, et al. Computational approaches for predicting the biological effect of p53 missense mutations: a comparison of three sequence analysis based methods. Nucleic Acids Res 2006;**34**(5):1317-25 doi: 10.1093/nar/gkj518[published Online First: Epub Date]|.
31. Tavtigian SV, Byrnes GB, Goldgar DE, et al. Classification of rare missense substitutions, using risk surfaces, with genetic- and molecular-epidemiology applications. Hum Mutat 2008;**29**(11):1342-54 doi: 10.1002/humu.20896[published Online First: Epub Date]|.
32. Vallee MP, Di Sera TL, Nix DA, et al. Adding In Silico Assessment of Potential Splice Aberration to the Integrated Evaluation of BRCA Gene Unclassified Variants. Hum Mutat 2016;**37**(7):627-39 doi: 10.1002/humu.22973[published Online First: Epub Date]|.
33. Thompson BA, Greenblatt MS, Vallee MP, et al. Calibration of multiple in silico tools for predicting pathogenicity of mismatch repair gene missense substitutions. Hum Mutat 2013;**34**(1):255-65 doi: 10.1002/humu.22214[published Online First: Epub Date]|.
34. Fortuno C, James PA, Young EL, et al. Improved, ACMG-Compliant, in silico prediction of pathogenicity for missense substitutions encoded by TP53 variants. Hum Mutat 2018 doi: 10.1002/humu.23553[published Online First: Epub Date]|.
35. Kim MP, Zhang Y, Lozano G. Mutant p53: Multiple Mechanisms Define Biologic Activity in Cancer. Front Oncol 2015;**5**:249 doi: 10.3389/fonc.2015.00249[published Online First: Epub Date]|.
36. Raue F, Frank-Raue K. Update multiple endocrine neoplasia type 2. Fam Cancer 2010;**9**(3):449-57 doi: 10.1007/s10689-010-9320-2[published Online First: Epub Date]|.
37. de la Hoya M, Soukarieh O, Lopez-Perolio I, et al. 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**(11):2256-68 doi: 10.1093/hmg/ddw094[published Online First: Epub Date]|.
38. Jager KJ, Zoccali C, Kramar R, et al. Measuring disease occurrence. Kidney Int 2007;**72**(4):412-5 doi: 10.1038/sj.ki.5002341[published Online First: Epub Date]|.
39. Thompson BA, Spurdle AB, Plazzer JP, et al. 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**(2):107-15 doi: 10.1038/ng.2854[published Online First: Epub Date]|.

40. Meeks HD, Song H, Michailidou K, et al. BRCA2 Polymorphic Stop Codon K3326X and the Risk of Breast, Prostate, and Ovarian Cancers. J Natl Cancer Inst 2016;**108**(2) doi: 10.1093/jnci/djv315[published Online First: Epub Date]|.
41. Arbustini E, Sgarella A, Ferrari A, et al. RE: BRCA2 Polymorphic Stop Codon K3326X and the Risk of Breast, Prostate, and Ovarian Cancers. J Natl Cancer Inst 2016;**108**(12) doi: 10.1093/jnci/djw172[published Online First: Epub Date]|.
42. Spurdle AB, Domchek S, Robson M, et al. Response. J Natl Cancer Inst 2016;**108**(12) doi: 10.1093/jnci/djw173[published Online First: Epub Date]|.
43. Lee A, Mavaddat N, Wilcox AN, et al. BOADICEA: a comprehensice breast cancer risk prediciton model incorporating genetic and non-genetic risk factors. Genetics in Medicine Provisionally accepted

Towards controlled terminology for reporting germline cancer susceptibility variants: an ENIGMA report.
Spurdle et al, J Medical Genetics.

Supplementary Table 1: Examples of variant descriptors and terms that have multiple meanings and/or appear to be used interchangeably.

Descriptor/s or Term/s that may be used interchangeably	**Alternative uses for descriptors or terms in literature** (Some specific source/s noted)
Mutation	● Permanent change in the nucleotide sequence [1]. ● Process by which a change occurs in DNA OR describing sequence changes in specific genes which are known to harbor disease-causing changes that result in well recognized heritable disorders (HVP, http://www.humanvariomeproject.org). ● Any germline change (relative to a given reference sequence), sometimes defined as "rare" with variable definitions of rare e.g. MAF <1%, <0.1%, <0.001%). ● Cancer-related change (*inferred* - relative to a given reference sequence or relative to the germline sequence for that individual) (CIVIC, https://civicdb.org/home). ● Germline sequence variant associated with phenotype of clinical relevance and variant carrier status may be used to inform clinical management of probands and/or relatives (NCCN, www.nccn.org; NICE, https://www.nice.org.uk; eviQ, https://www.eviq.org.au; CIMBA, http://cimba.ccge.medschl.cam.ac.uk).
Alteration	● Any germline change (relative to reference sequence). ● Any somatic change (relative to reference or germline sequence). ● Copy number variant or large genomic rearrangement.
Pathogenic Variant*; High-risk/High Penetrance Variant; Disease-predisposing	● Variant associated with/causing dominantly inherited disease. ● Health-care provider can use the molecular testing information in clinical decision-making (**Pathogenic** [1]). ● Variant associated with/causing phenotype of clinical relevance; variant carrier status may be used to inform clinical management of probands and/or relatives (**Pathogenic**, IARC scheme [2]). ● Variant contributes mechanistically to disease but is not necessarily fully penetrant i.e. may not be sufficient in isolation to cause disease (**Pathogenic** [3]).

Variant	● Variant significantly associated with increased risk (compared to general population average lifetime risk) of a specific disease or diseases, where the relevance of this increased risk to individual clinical management is unclear, or may differ depending on patient presentation.
Deleterious/ Damaging/ Possibly Damaging/ Disease-causing Variant Impacts Function/ Non-functional	● Variant reduces the reproductive fitness of carriers, and would thus be targeted by purifying natural selection (**Deleterious** [3]). ● Variant assessed by evolutionary/phylogenetic methods that is conserved in evolutionary space ~ evolutionarily unlikely change (**Damaging**, PolyPhen-2 (http://genetics.bwh.harvard.edu/pph2) & SIFT (http://sift-dna.org), See Table 1). ● Descriptor for variant predicted bioinformatically to alter protein structure or function (**Damaging**, FATHMM, See Table 1). ● Variant that causes loss of protein function, either by impacting mRNA level or transcript profile and/or by resulting in a change in encoded protein sequence that affects protein stability, conformation or function (**Damaging** – for loss of protein function [4]). ● Sequence variant associated with phenotype of clinical relevance and variant carrier status may be used alone to inform clinical management of probands and/or relatives (**Deleterious**, NCCN (www.nccn.org), EMA (http://www.ema.europa.eu/), FDA (https://www.fda.gov/)). ● Descriptor for variant that abrogates or alters the normal levels or biochemical function of a gene or gene product, generally determined using *in vitro* or *in vivo* assays (**Damaging** [3 4], **Deleterious** [5 6], **Impacts function** [7 8].
Null Variant; Loss of Function (LOF) Variant Non-functional Variant	● Nonsense, frameshift, canonical +1/2, initiation codon and single exon deletion, multi-exon deletion variant (**Null** [1]). ● As above, but with multiple other gene-specific considerations and additional details that infer loss of function without experimental evidence (**LOF**, ClinGen Sequence Variant Interpretation Committee, [9]). ● Variant proven from molecular assays to result in loss of function at the RNA or protein level – due to effect of mRNA transcription or stability, or abrogated protein function (**non-functional**, [10]).
Moderate risk/ Intermediate risk Variant	● Germline variant that is significantly associated with risk of disease at a level (generally >2-fold) that may be used to alter clinical management in selected circumstances or in consideration with other risk factors e.g. prevention for second cancers, for individuals with family history of the disease or other risk modifiers (**Moderate** [11]). ● Germline variant that is significantly associated with risk of disease that is intermediate between general

	population risk and that of a classical high-risk pathogenic variant in that gene (**Intermediate** [12 13]; **Moderate** [14]).
Polymorphism; Common Variant; SNP, SNV	● A common sequence change (frequency not defined) ● Germline sequence change not associated with a phenotype of clinical relevance, irrespective of variant frequency. ● Germline sequence change associated with risk of disease, at a level that is not considered clinically relevant, on its own, to alter patient management recommendations. ● Variant with a population frequency above 1% (**Polymorphism** [1]) ● Variant assessed by bioinformatic methods that include evolutionary/phylogenetic methods and protein structure/function and considered unlikely to alter function/cause disease, irrespective of variant frequency (**Polymorphism**, LRT bioinformatics tool, http://www.genetics.wustl.edu/jflab/lrt_query.html. See Table 1) ● Single nucleotide polymorphism (SNP) ~ germline single nucleotide substitution sequence variant (frequency may not be defined) ● Single nucleotide variant (SNV) ~ commonly used in the context of high throughput sequencing to describe any single nucleotide substitution sequence variant (germline or tumor) ● Phenomenon in biology, describing discontinuous genetic variation resulting in the occurrence of several different forms (**Genetic Polymorphism** [15])
Neutral/Tolerated Variant	● Variant not associated with a phenotype of clinical relevance (**Neutral**, initial integrated evaluation publication [16], also [5 7 8]. ● Variant assessed by evolutionary/phylogenetic methods and not conserved (**Tolerated** or **Neutral,** multiple bioinformatic tools, See Table 1). ● Variant that does not cause loss of protein function, either by impacting mRNA level or transcript profile and/or by resulting in a change in encoded protein sequence that affects protein stability, conformation or function (**Tolerated**, **evolutionarily neutral** [17]).
Benign/Innocuous/ Harmless Variant	● Variant not associated with a phenotype of clinical relevance (**Benign**, ACMG/AMP, [1]). ● Germline change that is significantly associated with a very modest risk of disease as measured using large-scale genetic studies e.g <1.5-fold or 2-fold, and is not used alone to inform clinical management . ● Variant that does not cause loss of protein function, either by impacting mRNA level or transcript profile and/or by resulting in a change in encoded protein sequence that affects protein stability, conformation or function. ● Variant assessed by bioinformatic methods and considered unlikely to alter function/cause disease (**Benign**,

	PolyPhen-2, http://genetics.bwh.harvard.edu/pph2. See Table 1).
Not pathogenic or non-pathogenic Variant/ Variant of no clinical significance/ Variant of little clinical significance	● Variant not associated with a phenotype of clinical relevance (**Not pathogenic/no clinical significance**, IARC scheme [2]). ● Germline variant that is significantly associated with a very modest risk of disease as measured using large-scale genetic studies e.g <1.5-fold, that is not used alone to inform clinical management (**Not pathogenic/no clinical significance** [18]).
Unclassified/ Unknown/Uncertain Variant; Variant of Uncertain Clinical Significance/ Variant of Uncertain Significance	● Variant that has not yet been assessed for its association with phenotype of clinical relevance (**Unclassified** [4]). ● Variant that has been assessed for its association with phenotype of clinical relevance but for which clinical significance remains uncertain (**Uncertain** [4]; **Unclassified** [19]).

Abbreviations: ACMG – American College of Medical Genetics and Genomics; AMP - Association for Molecular Pathology; CIMBA - Consortium of Investigators of Modifiers of *BRCA1* and *BRCA2*; CIViC - Clinical Interpretation of Variants in Cancer; EMA - European Medicines Agency; FDA - US Food & Drug administration; HVP – Human Variome Project; LRT – Likelihood Ratio Test; MAF - minor allele frequency; NCCN - National Comprehensive Cancer Network; NICE - National Institute for Health and Care Excellence; IARC – International Agency for Research into Cancer.

** Note, a sequence variant implicated in risk of another (recessive) disease may be termed pathogenic without considering its relevance to the disease being evaluated. Also see Table 4. .*
*** In some circumstances a low-risk variant may be misconstrued to be a variant that lowers risk of disease (as opposed to a variant that is associated with a low but measurable increased risk of disease)*

References:

1. Richards S, Aziz N, Bale S, et al. 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**(5):405-24 doi: 10.1038/gim.2015.30[published Online First: Epub Date]|.
2. Plon SE, Eccles DM, Easton D, et al. Sequence variant classification and reporting: recommendations for improving the interpretation of cancer susceptibility genetic test results. Hum Mutat 2008;**29**(11):1282-91 doi: 10.1002/humu.20880[published Online First: Epub Date]|.
3. MacArthur DG, Manolio TA, Dimmock DP, et al. Guidelines for investigating causality of sequence variants in human disease. Nature 2014;**508**(7497):469-76 doi: 10.1038/nature13127[published Online First: Epub Date]|.
4. Lindor NM, Guidugli L, Wang X, et al. A review of a multifactorial probability-based model for classification of BRCA1 and BRCA2 variants of uncertain significance (VUS). Hum Mutat 2012;**33**(1):8-21 doi: 10.1002/humu.21627[published Online First: Epub Date]|.
5. Ransburgh DJ, Chiba N, Ishioka C, et al. Identification of breast tumor mutations in BRCA1 that abolish its function in homologous DNA recombination. Cancer Res 2010;**70**(3):988-95 doi: 10.1158/0008-5472.CAN-09-2850[published Online First: Epub Date]|.
6. Thouvenot P, Ben Yamin B, Fourriere L, et al. Functional Assessment of Genetic Variants with Outcomes Adapted to Clinical Decision-Making. PLoS Genet 2016;**12**(6):e1006096 doi: 10.1371/journal.pgen.1006096[published Online First: Epub Date]|.
7. Guidugli L, Carreira A, Caputo SM, et al. Functional assays for analysis of variants of uncertain significance in BRCA2. Hum Mutat 2014;**35**(2):151-64 doi: 10.1002/humu.22478[published Online First: Epub Date]|.
8. Millot GA, Carvalho MA, Caputo SM, et al. A guide for functional analysis of BRCA1 variants of uncertain significance. Hum Mutat 2012;**33**(11):1526-37 doi: 10.1002/humu.22150[published Online First: Epub Date]|.
9. Abou Tayoun AN, Pesaran T, DiStefano MT, et al. Recommendations for interpreting the loss of function PVS1 ACMG/AMP variant criterion. Hum Mutat 2018 doi: 10.1002/humu.23626[published Online First: Epub Date]|.
10. Findlay GM, Daza RM, Martin B, et al. Accurate classification of BRCA1 variants with saturation genome editing. Nature 2018;**562**(7726):217-22 doi: 10.1038/s41586-018-0461-z[published Online First: Epub Date]|.
11. Easton DF, Pharoah PD, Antoniou AC, et al. Gene-panel sequencing and the prediction of breast-cancer risk. N Engl J Med 2015;**372**(23):2243-57 doi: 10.1056/NEJMsr1501341[published Online First: Epub Date]|.
12. Spurdle AB, Whiley PJ, Thompson B, et al. BRCA1 R1699Q variant displaying ambiguous functional abrogation confers intermediate breast and ovarian cancer risk. J Med Genet 2012;**49**(8):525-32 doi: 10.1136/jmedgenet-2012-101037[published Online First: Epub Date]|.
13. Moghadasi S, Meeks HD, Vreeswijk MP, et al. 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**(1):15-20 doi: 10.1136/jmedgenet-2017-104560[published Online First: Epub Date]|.

14. Shimelis H, Mesman RLS, Von Nicolai C, et al. BRCA2 Hypomorphic Missense Variants Confer Moderate Risks of Breast Cancer. Cancer Res 2017;**77**(11):2789-99 doi: 10.1158/0008-5472.CAN-16-2568[published Online First: Epub Date]|.
15. Ford EB. *Genetic Polymorphism*. London: Faber & Faber, 1965.
16. Goldgar DE, Easton DF, Deffenbaugh AM, et al. Integrated evaluation of DNA sequence variants of unknown clinical significance: application to BRCA1 and BRCA2. Am J Hum Genet 2004;**75**(4):535-44 doi: 10.1086/424388[published Online First: Epub Date]|.
17. Hietpas RT, Jensen JD, Bolon DN. Experimental illumination of a fitness landscape. Proc Natl Acad Sci U S A 2011;**108**(19):7896-901 doi: 10.1073/pnas.1016024108[published Online First: Epub Date]|.
18. Spurdle AB, Domchek S, Robson M, et al. Response. J Natl Cancer Inst 2016;**108**(12) doi: 10.1093/jnci/djw173[published Online First: Epub Date]|.
19. Mur P, Sanchez-Cuartielles E, Ausso S, et al. Scarce evidence of the causal role of germline mutations in UNC5C in hereditary colorectal cancer and polyposis. Sci Rep 2016;**6**:20697 doi: 10.1038/srep20697[published Online First: Epub Date]|.

Towards controlled terminology for reporting germline cancer susceptibility variants: an ENIGMA report. Spurdle et al, J Medical Genetics.

Supplementary Table 2: Examples of multi-tier reports for variants potentially identified by multi-gene cancer panel testing (Relevant genes for example variant reports are *ATM* (MIM607585), *BRCA1* (MIM113705), *BRCA2* (MIM600185), *CHEK2* (604373).

Demographic Information	Female, age 55, no family history
Cancer Phenotype	Breast Cancer
Sample Tissue	√ Blood
Context	√ Diagnostic √ Prediction Cancer Risk √ Genotype Directed Treatment x Other (specify)

Variant Identified	NM_007294.3(BRCA1):c.1969C>T p.(Gln657Ter)

Level 1: Variant Classification	
IARC Variant Classification	**Pathogenic**. IARC guidelines for the interpretation and classification of variants present strong evidence that this variant is or may be causal for one or more cancer phenotypes including breast and ovarian cancer. Details of the evidence supporting the variant classification is provided as supplementary documentation.

Level 2: Clinical Validity	
Presenting Cancer Phenotype	There is strong evidence that this variant is making a substantial contribution to the presenting cancer phenotype [1].
Prediction Cancer Risk	There is strong evidence to support the prediction of a high absolute breast and ovarian cancer risk when this variant is present in a family member. Based upon prospective study data, the average cumulative lifetime risk is 60% (range 44%-75%) and the average cumulative lifetime risk for ovarian cancer risk is 59% (43-76%) [2 3].
Genotype Directed Treatment	There is evidence to support consideration of PARP inhibitors in the context of breast and ovarian cancer [4].
Biallelic Inheritance	Rare reported cases of biallelic (compound heterozygote or homozygote) inheritance of *BRCA1* pathogenic variants have been reported to cause Fanconi Anemia [5-8]

Level 3: Clinical Utility & Actionability	
Action	This variant is clinically actionable. Refer to clinical genetics service for full evaluation and discussion of risk management options. 1. High risk breast surveillance. 2. Eligible to consider risk reducing surgery depending on cancer prognosis and treatment. This may include consideration of BRRM [9]. BSO likely to be advised once childbearing complete. 3. Consideration of precision management with PARP inhibitors and platinum based chemotherapy. 4. Cascade genetic testing

Demographic Information	Female age 30 years, benign breast lump investigated. Concerned about family history
Cancer Phenotype	Unaffected
Sample Tissue	√ Blood
Context	x Diagnostic √ Prediction Cancer Risk x Genotype Directed Treatment x Other (specify)

Variant Identified	NM_000059.3(BRCA2):c.5delC p.(Pro2LeufsTer23)

Level 1: Variant Classification	
IARC Variant Classification	**Pathogenic.** IARC guidelines for the interpretation and classification of variants present strong evidence that this variant is or may be causal for one or more cancer phenotypes including breast and ovarian cancer. Details of the evidence supporting the variant classification is provided as supplementary documentation.

Level 2: Clinical Validity	
Presenting Cancer Phenotype	N/A
Prediction Cancer Risk	There is strong evidence to support the prediction of a high absolute breast and ovarian cancer risk when this variant is present in an unaffected female. Based upon prospective study data, the average cumulative lifetime risk for breast cancer is 55% (range 41-72%) and for ovarian cancer the average cumulative lifetime risk is 16.5% (range 7.5-34%) [2 3].
Genotype Directed Treatment	N/A
Biallelic Inheritance	Evidence supports that biallelic (compound heterozygote or homozygote) variant inheritance causes Fanconi anaemia [10]

Level 3: Clinical Utility & Actionability	
Action	This variant is clinically actionable. Refer to clinical genetics service for full evaluation and discussion of risk management options. 1. High risk breast surveillance. 2. Eligible to consider risk reducing surgery depending on cancer prognosis and treatment. This may include consideration of BRRM [9]. BSO likely to be advised once childbearing complete. 3. Consideration of precision management with PARP inhibitors and platinum based chemotherapy if develops cancer. 4. Cascade genetic testing

Demographic Information	Female aged 43
Cancer Phenotype	Breast Cancer
Sample Tissue	√ Blood
Context	√ Diagnostic √ Prediction Cancer Risk √ Genotype Directed Treatment x Other (specify)

Variant Identified	NM_007294.3(BRCA1):c.5096G>A p.(Arg1699Gln)

Level 1: Variant Classification	
ACMG Variant Classification	**Pathogenic**. ACMG guidelines for the interpretation and classification of variants present strong evidence that this variant is or may be contributing to risk of one or more cancer phenotypes including breast cancer. Details of the evidence supporting the variant classification is provided as supplementary documentation.

Level 2: Clinical Validity	
Presenting Cancer Phenotype	There is strong evidence that this variant is contributing to the presenting cancer phenotype [11].
Prediction Cancer Risk	There is strong evidence to support a moderate increase in new primary breast and ovarian cancer risk when this variant is present in an unaffected family member and can be used for cancer risk stratification [12 13].
Genotype Directed Treatment	There is no currently evidence that reduced penetrance BRCA1 variants are associated with sensitivity to targeted treatment with PARP inhibitors.
Biallelic Inheritance	There are no reported cases of Fanconi anaemia due to biallelic (compound heterozygote or homozygote) with this particular variant but it remains hypothetically possible.

Level 3: Clinical Utility & Actionability	
Action	This variant is clinically actionable. Refer to clinical genetics service for full evaluation and discussion of risk management options. 1. Carriers should be offered moderate risk surveillance 2. Level of risk is not sufficiently high to be eligible for risk reducing breast surgery (estimate 20% by age 70) 3. Lifetime ovarian cancer risk is 6% and female carriers of this variant could consider risk reducing salpingo-oophorectomy 4. Insufficient evidence regarding the clinical efficacy of PARP inhibitors but platinum based chemotherapy remains a consideration. 5. Cascade genetic testing

Demographic Information	Male, aged 45
Cancer Phenotype	Colorectal Cancer
Sample Tissue	√ Blood
Context	√ Diagnostic √ Prediction Cancer Risk √ Genotype Directed Treatment x Other (specify)

Variant Identified	NM_000051.3 (ATM): c.7271T>G p.(Val2424Gly)

Level 1: Variant Classification	
ACMG Variant Classification	**Pathogenic.** ACMG guidelines for the interpretation and classification of variants present strong evidence that this variant is or may contribute substantially to one or more cancer phenotypes including breast cancer. Details of the evidence supporting the variant classification is provided as supplementary documentation.

Level 2: Clinical Validity	
Presenting Cancer Phenotype	There is insufficient evidence to support an association between this variant and the presenting colorectal cancer phenotype.
Prediction Cancer Risk	There is strong evidence to support the prediction of a high absolute risk of new primary breast cancer when this variant is present in a female family member. This magnitude of effect can be in excess of 8-11 fold compared to the basal population risk [14 15].
Genotype Directed Treatment	There is currently no evidence to support consideration of precision treatment for pathogenic *ATM* variant carriers in the context of colorectal cancer, but they may be eligible for clinical trials.
Biallelic Inheritance	Evidence may support that biallelic (compound heterozygote or homozygote) variant inheritance is likely to cause Ataxia Telangiectasia [16].

Level 3: Clinical Utility & Actionability	
Action	This variant is clinically actionable. Refer to clinical genetics service for full evaluation and discussion of risk management options. 1. High risk breast surveillance in female carriers 2. Cascade genetic testing 3. Bowel cancer risk should be managed according to the family history, with referral to a regional family history service following local guidelines if required.

Demographic Information	Female, aged 48
Cancer Phenotype	Breast Cancer
Sample Tissue	√ Blood
Context	√ Diagnostic √ Prediction Cancer Risk √ Genotype Directed Treatment x Other (specify)

Variant Identified	NM_000051.3 (ATM): c.2413C>T p.(Arg805Ter)

Level 1: Variant Classification	
ACMG Variant Classification	**Pathogenic.** ACMG guidelines for the interpretation and classification of variants present strong evidence that this variant is or may be contributing to breast cancer phenotype. Details of the evidence supporting the variant classification is provided as supplementary documentation.

Level 2: Clinical Validity	
Presenting Cancer Phenotype	There is evidence to support an association between this variant and the presenting cancer phenotype [17].
Prediction Cancer Risk	There is strong evidence to support the prediction of a ~3-fold increased risk of breast cancer, compared to the basal population risk in female gene carriers [17]. 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 further modify overall cancer risk estimates.
Genotype Directed Treatment	There is currently no evidence to support use of pathogenic *ATM* variant carrier status to predict response to specific drug treatments.
Biallelic Inheritance	Evidence supports biallelic (compound heterozygote or homozygote) variant inheritance is likely to cause Ataxia Telangiectasia [16].

Level 3: Clinical Utility & Actionability	
Action	This variant is clinically actionable but not in isolation. Refer to clinical genetics service and local guidelines where available. A full risk evaluation incorporating family history and other known risk factors should inform discussion of risk management options+. 1. Additional risk surveillance may be required. 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. The level of risk for this variant alone is not sufficiently high to advocate risk reducing breast surgery. 3. Consideration of cascade genetic testing based upon local (regional/national) guidance.
+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.	

Demographic Information	Male aged 54, mother died of breast cancer age 52, two daughters,
Cancer Phenotype	N/A
Sample Tissue	√ Blood
Context	x Diagnostic √ Prediction Cancer Risk x Genotype Directed Treatment x Other (specify)

Variant Identified	NM_007194.3 (CHEK2): c.1100delC p.(Thr367Metfs)

Level 1: Variant Classification	
ACMG Variant Classification	**Pathogenic**. ACMG guidelines for the interpretation and classification of variants present strong evidence that this variant is or may be contributing to one or more cancer phenotypes including breast cancer. Details of the evidence supporting the variant classification is provided as supplementary documentation.

Level 2: Clinical Validity	
Presenting Cancer Phenotype	N/A
Prediction Cancer Risk	There is strong evidence to support a moderate increase in breast cancer risk associated with *CHEK2* c.1100delC p.(Thr367Metfs) when this variant is present in a family member, and it can be used for cancer risk stratification. The risk is generally considered to be 2-3 fold above baseline population risk [18 19]. 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 further modify overall cancer risk estimates [20].
Genotype Directed Treatment	N/A
Biallelic Inheritance	Biallelic mutation carriers have a high absolute lifetime risk of breast cancer [21] .

Level 3: Clinical Utility & Actionability	
Action	This variant is clinically actionable but not in isolation. Refer to clinical genetics service and local guidelines where available. A full risk evaluation incorporating family history and other known risk factors should inform discussion of risk management options+. 1. Additional risk surveillance may be required in female carriers. 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. The level of risk for this variant alone is not sufficiently high to advocate risk reducing breast surgery. 3. Consideration of cascade genetic testing based on local (regional/national) guidance.
+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.	

Demographic Information	Female aged 40, twins by IVF age 2 years.
Cancer Phenotype	N/A
Sample Tissue	√ Blood
Context	x Diagnostic √ Prediction Cancer Risk x Genotype Directed Treatment x Other (specify)

Variant Identified	NM_007194.3 (CHEK2): c.470T>C p.(Ile157Thr)

Level 1: Variant Classification	
ACMG Variant Classification	**Little clinical significance / Benign**. Population studies show this sequence variant is associated with a low increased risk of cancer [22 23]. Details of the evidence supporting the variant classification is provided as supplementary documentation.

Level 2: Clinical Validity	
Presenting Cancer Phenotype	N/A
Prediction Cancer Risk	There is evidence from large population studies to support a low increase in breast cancer risk, but the variant cannot be used in isolation for cancer risk prediction. The frequency of this variant is significantly higher amongst unselected female breast cancer patients compared with the expected population frequencies with an estimated magnitude of breast cancer risk which is approximately 1.5 fold compared to the basal population level [22 23]. This variant alone is unhelpful in predicting cancer risk for an individual.
Genotype Directed Treatment	N/A
Biallelic Inheritance	There are no data available regarding the impact of biallelic inheritance of this specific variant on cancer risk or risk for other phenotypes.

Level 3: Clinical Utility & Actionability	
Action	This variant is not clinically actionable. 1. Manage Based Upon Family History. This variant is insufficiently predictive of individual cancer risk to be clinically actionable. Clinical management recommendations should be determined on the basis of personal and/or family history of disease and other known risk factors. Variants in this category may contribute towards a polygenic risk score.

References:

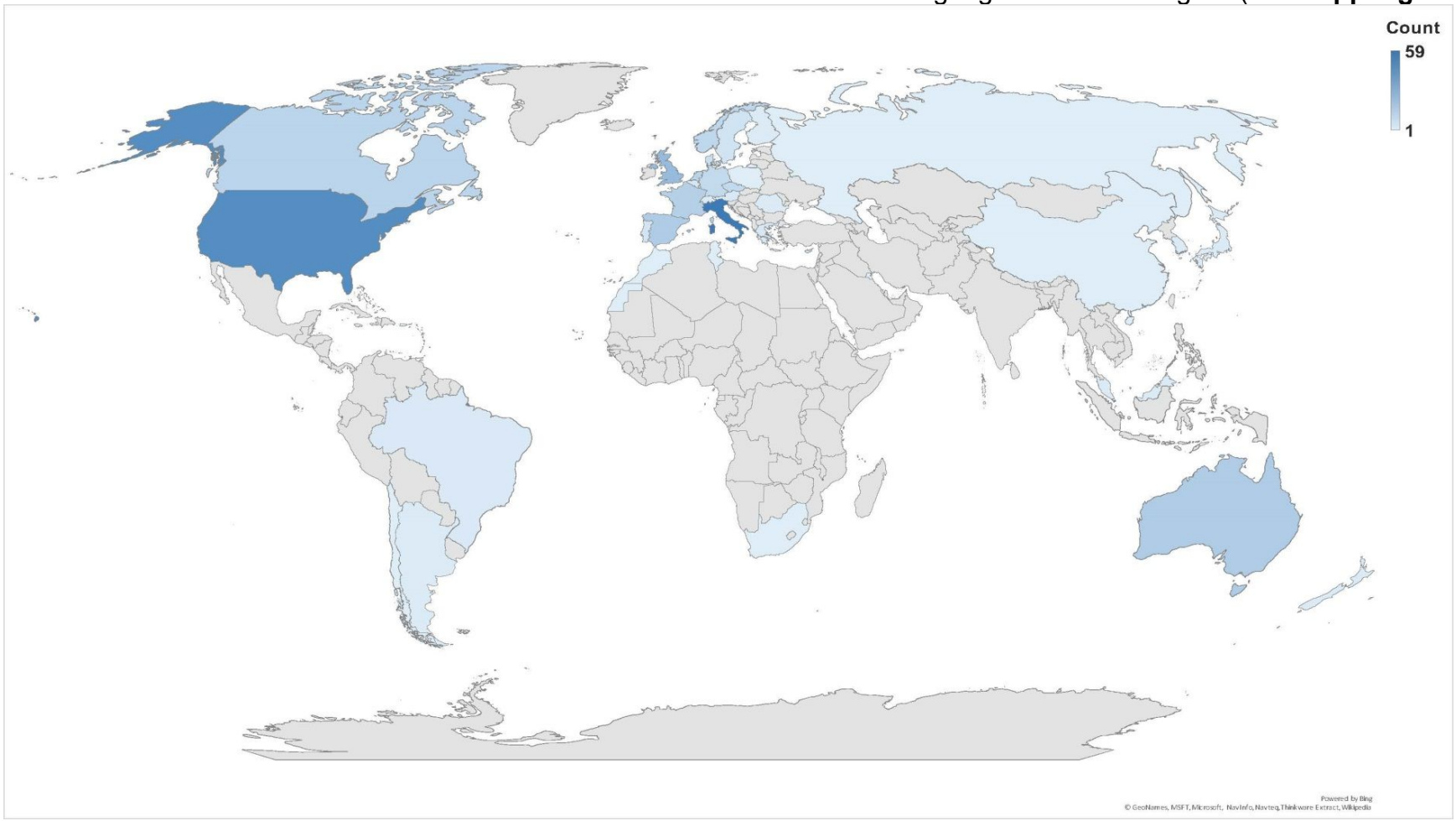
1. Miki Y, Swensen J, Shattuck-Eidens D, et al. A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. Science 1994;**266**(5182):66-71
2. Mavaddat N, Peock S, Frost D, et al. Cancer Risks for BRCA1 and BRCA2 Mutation Carriers: Results From Prospective Analysis of EMBRACE. JNCI: Journal of the National Cancer Institute 2013;**105**(11):812-22 doi: 10.1093/jnci/djt095[published Online First: Epub Date]|.
3. Kuchenbaecker KB, Hopper JL, Barnes DR, et al. Risks of Breast, Ovarian, and Contralateral Breast Cancer for BRCA1 and BRCA2 Mutation Carriers. JAMA 2017;**317**(23):2402-16 doi: 10.1001/jama.2017.7112[published Online First: Epub Date]|.
4. Robson M, Im S-A, Senkus E, et al. Olaparib for Metastatic Breast Cancer in Patients with a Germline BRCA Mutation. New England Journal of Medicine 2017;**377**(6):523-33 doi: 10.1056/NEJMoa1706450[published Online First: Epub Date]|.
5. Domchek SM, Tang J, Stopfer J, et al. Biallelic deleterious BRCA1 mutations in a woman with early-onset ovarian cancer. Cancer Discov 2013;**3**(4):399-405 doi: 10.1158/2159-8290.CD-12-0421[published Online First: Epub Date]|.
6. Sawyer SL, Tian L, Kahkonen M, et al. Biallelic mutations in BRCA1 cause a new Fanconi anemia subtype. Cancer Discov 2015;**5**(2):135-42 doi: 10.1158/2159-8290.CD-14-1156[published Online First: Epub Date]|.
7. Freire BL, Homma TK, Funari MFA, et al. Homozygous loss of function BRCA1 variant causing a Fanconi-anemia-like phenotype, a clinical report and review of previous patients. Eur J Med Genet 2018;**61**(3):130-33 doi: 10.1016/j.ejmg.2017.11.003[published Online First: Epub Date]|.
8. Seo A, Steinberg-Shemer O, Unal S, et al. Mechanism for survival of homozygous nonsense mutations in the tumor suppressor gene BRCA1. Proc Natl Acad Sci U S A 2018;**115**(20):5241-46 doi: 10.1073/pnas.1801796115[published Online First: Epub Date]|.
9. Hartmann LC, Lindor NM. The Role of Risk-Reducing Surgery in Hereditary Breast and Ovarian Cancer. New England Journal of Medicine 2016;**374**(5):454-68 doi: 10.1056/NEJMra1503523[published Online First: Epub Date]|.
10. Meyer S, Tischkowitz M, Chandler K, et al. Fanconi anaemia, BRCA2 mutations and childhood cancer: a developmental perspective from clinical and epidemiological observations with implications for genetic counselling. J Med Genet 2014;**51**(2):71-5 doi: 10.1136/jmedgenet-2013-101642[published Online First: Epub Date]|.
11. Spurdle AB, Whiley PJ, Thompson B, et al. BRCA1 R1699Q variant displaying ambiguous functional abrogation confers intermediate breast and ovarian cancer risk. J Med Genet 2012;**49**(8):525-32 doi: 10.1136/jmedgenet-2012-101037[published Online First: Epub Date]|.
12. Shimelis H, Mesman RLS, Von Nicolai C, et al. BRCA2 Hypomorphic Missense Variants Confer Moderate Risks of Breast Cancer. Cancer Res 2017;**77**(11):2789-99 doi: 10.1158/0008-5472.CAN-16-2568[published Online First: Epub Date]|.
13. Moghadasi S, Meeks HD, Vreeswijk MP, et al. 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**(1):15-20 doi: 10.1136/jmedgenet-2017-104560[published Online First: Epub Date]|.
14. Tavtigian SV, Oefner PJ, Babikyan D, et al. Rare, evolutionarily unlikely missense substitutions in ATM confer increased risk of breast cancer. Am J Hum Genet 2009;**85**(4):427-46 doi: 10.1016/j.ajhg.2009.08.018[published Online First: Epub Date]|.
15. Southey MC, Goldgar DE, Winqvist R, et al. PALB2, CHEK2 and ATM rare variants and cancer risk: data from COGS. J Med Genet 2016;**53**(12):800-11 doi: 10.1136/jmedgenet-2016-103839[published Online First: Epub Date]|.

16. Renwick A, Thompson D, Seal S, et al. ATM mutations that cause ataxia-telangiectasia are breast cancer susceptibility alleles. Nat Genet 2006;**38**(8):873-5 doi: 10.1038/ng1837[published Online First: Epub Date]|.
17. van Os NJ, Roeleveld N, Weemaes CM, et al. Health risks for ataxia-telangiectasia mutated heterozygotes: a systematic review, meta-analysis and evidence-based guideline. Clin Genet 2016;**90**(2):105-17 doi: 10.1111/cge.12710[published Online First: Epub Date]|.
18. Decker B, Allen J, Luccarini C, et al. Rare, protein-truncating variants in ATM, CHEK2 and PALB2, but not XRCC2, are associated with increased breast cancer risks. J Med Genet 2017;**54**(11):732-41 doi: 10.1136/jmedgenet-2017-104588[published Online First: Epub Date]|.
19. Schmidt MK, Hogervorst F, van Hien R, et al. Age- and Tumor Subtype-Specific Breast Cancer Risk Estimates for CHEK2*1100delC Carriers. J Clin Oncol 2016;**34**(23):2750-60 doi: 10.1200/JCO.2016.66.5844[published Online First: Epub Date]|.
20. Rudolph A, Song M, Brook MN, et al. Joint associations of a polygenic risk score and environmental risk factors for breast cancer in the Breast Cancer Association Consortium. Int J Epidemiol 2018;**47**(2):526-36 doi: 10.1093/ije/dyx242[published Online First: Epub Date]|.
21. Adank MA, Jonker MA, Kluijt I, et al. CHEK2*1100delC homozygosity is associated with a high breast cancer risk in women. J Med Genet 2011;**48**(12):860-3 doi: 10.1136/jmedgenet-2011-100380[published Online First: Epub Date]|.
22. Kilpivaara O, Vahteristo P, Falck J, et al. CHEK2 variant I157T may be associated with increased breast cancer risk. Int J Cancer 2004;**111**(4):543-7 doi: 10.1002/ijc.20299[published Online First: Epub Date]|.
23. Cybulski C, Wokolorczyk D, Jakubowska A, et al. Risk of breast cancer in women with a CHEK2 mutation with and without a family history of breast cancer. J Clin Oncol 2011;**29**(28):3747-52 doi: 10.1200/JCO.2010.34.0778[published Online First: Epub Date]|.

SUPPLEMENTARY TEXT: Towards controlled terminology for reporting germline cancer susceptibility variants: an ENIGMA report. Spurdle et al, J Medical Genetics.

ENIGMA membership

Eligibility for membership is broad: an ENIGMA member is currently defined as a researcher or research group (consortium) who is willing to work collaboratively towards classification of variants by contributing data from families and/or conducting statistical analysis or laboratory-based assays within a working group framework, namely clinical, analytical, functional, splicing and pathology working groups. The ENIGMA membership listing currently includes 309 individuals from 38 different countries. Most of these members (202) are from countries where research and clinical activities are conducted in a language other than English (see **Supp Figure 1**).



Supplementary Figure 1: Heatmap showing distribution of ENIGMA members across countries.

Published recommendations on variant terminology

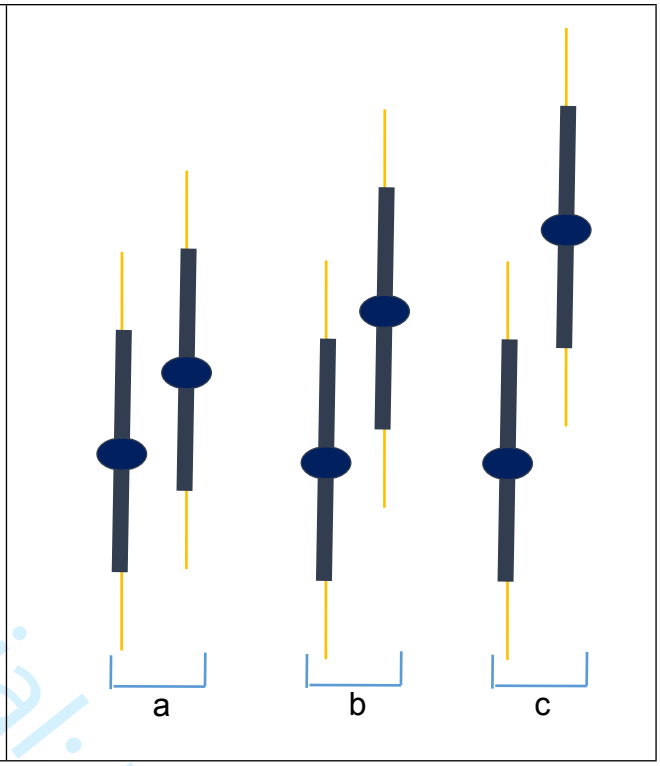
ACMG/AMP [1]: "The terms "mutation" and "polymorphism," however, which have been used widely, often lead to confusion because of incorrect assumptions of pathogenic and benign effects, respectively. Thus, it is recommended that both terms be replaced by the term "variant" with the following modifiers: (i) pathogenic, (ii) likely pathogenic, (iii) uncertain significance, (iv) likely benign, or (v) benign."
HVP [2]: "The term "variant" should be used to describe all sequences changes irrespective of their contribution to phenotype. Mutation may be used in the correct sense of the word to describe the process by which variants arise. Use of the term polymorphism is deprecated."
It is also the term used in publication to describe somatic alterations by the Variant Interpretation for Cancer Consortium (**VICC**) [3], although the terminology seems not to be transferred to the CIViC web-based resource for expert crowdsourcing of Clinical Interpretation of Variants in Cancer (http://civicdb.org/), a key output from this group.

Clinical calibration, replication and validation

Clinical calibration refers to the process where a specific type of information (e.g. bioinformatic score, protein function, splicing aberration, pathology information) is calibrated as a measure of variant pathogenicity against clinical predictors of variant pathogenicity (e.g. segregation data, family history profile, frequency in population controls). For example, Easton et al [4] provides an assessment of family history profiles of *BRCA1/2* pathogenic variant carriers versus non-carriers, according to variant location in specific motifs or domains, combined with bioinformatic prediction of missense effect. To avoid over-fitting, we recommend that initial calibration of a promising predictive measure (in particular functional assays) should be followed by a validation study using an independent set of variants, or at least an independent patient observational dataset. Further, as shown in **Supplementary Figure 2** below, and utilized in Drost *et al* [5], we suggest the following approach to describe replication and validation of a clinical calibration exercise. A functional assay method and its calibration can be considered to be replicated, *but not validated*, if at least one point estimated from the Training and Validation sets falls outside the other point estimate's 80% Confidence Interval (CI), but the point estimates are mutually within their 95% CI, which are wider. In this case, the method's point estimates and confidence intervals will be refined by recalculation from the combined data set; in this scenario it may be considered appropriate to withhold the method from quantitative clinical use until it has met the stricter criterion for replication and validation against a later data set. A method and its calibration are considered invalidated if either or both point estimates fall outside the other point estimate's 95% CI.

Supplementary Figure 2:
Depiction of confidence interval-based assay validation logic.

Blue ball: point estimate.
Blue line: 80% confidence limits.
Yellow line: 95% confidence limits.
a. Replicated **and** validated.
b. Replicated but **not** validated.
c. Invalidation.



Definitions of key terms used to describe the clinical importance of sequence variation.

Definitions, including health-related definitions where available, were sourced by online searches of the Oxford, Collins and Merriam-Webster dictionaries. Derivations were synthesized from all three sources. Considerations about translations of these terms into other languages was discussed by ENIGMA members whose primary language was not English.

Pathogenic:
Oxford: (of a bacterium, virus, or other microorganism) causing disease.
Collins: able to cause or produce disease (British), producing disease (American).
Merriam-Webster: Causing or capable of causing disease.
Derivation: From the Greek *pathos* (suffering) + *gen* (that which produces).

Deleterious:
Oxford: causing harm or damage
Collins: harmful, injurious, hurtful (English), harmful to health or well-being, injurious (American).
Merriam-Webster: damaging or harmful, or harmful often in a subtle or unexpected way (medical definition).
Derivation: via Medieval Latin from Greek *deleterios* (noxious/injurious/destructive).

Benign:
Oxford: (of a disease) not harmful in effect; (of a tumour) not malignant.
Collins: not threatening to life or health; not malignant (British), doing little or no harm, not malignant (American).
Merriam-Webster: of a mild type or character that does not threaten health or life (especially, not becoming cancerous); having no significant effect.
Derivation: Middle English - from Old French *benigne*, from Latin *benignus*, (probably) from *bene* (well) + *-genus* (-born). Alternative derivation (Collins) - from Old French *benigne*, from Latin *benignus*, from *bene* (well) + *gignere* (to produce).

Mutation:
Oxford: the process or an instance of change or alteration; a genetic change which, when transmitted to offspring, gives rise to heritable variations (Australian), action or process of mutating, changing of the structure of a gene – resulting in a variant form that may be transmitted to subsequent generations; a distinct form resulting from genetic mutation (English).
Collins: the act or process of mutating, change, alteration; a change or alteration; a change in the chromosomes or genes of a cell, which when in gametes the structure and development of the resultant offspring may be affected (English). a changing or being changed; a change as in form, nature, qualities; a sudden variation in some inheritable characteristic in a germ cell of an individual animal or plant, as distinguished from a variation resulting from generations of gradual change; an individual resulting from such variation; an abrupt and relatively permanent change in somatic cells that is transmitted only to daughter cell and can be inherited only in plants that reproduce asexually (American).
Merriam-Webster: a significant and basic alteration; a relatively permanent change in hereditary material that involves either a change in chromosome structure or number or a change in the nucleotide sequence of a gene's codons (as in frameshift or missense errors) and that occurs either in germ cells or in somatic cells but with only those in germ cells being capable of perpetuation by sexual reproduction; an individual, strain, or trait resulting from mutation
Derivation: Middle English from Latin *mutatio*, from *mutare* (change).

Significance:
Oxford: the quality of being worthy of attention, important.
Collins: consequence or importance.
Merriam-Webster: quality of being important, the quality of having notable worth or influence.
Derivation: Middle English via Old French *significance,* from Latin *significantia*, from *significare* (indicate, portend).

Annotation based on variant sequence/position and (predicted) effect on gene/protein product

Variant annotation, often used in the context of next generation sequencing bioinformatic pipelines, is the process of assigning various descriptors or data points relevant for a given sequence variant. The purpose is to assess variously: the quality of a sequence variant call; the location of a variant in relation to functional genomic regions (exons, introns, splice sites, regulatory sites); bioinformatically predicted effect on mRNA transcript/encoded protein, nucleotide protein or evolutionary conservation; and multiple other possible features used to interpret the clinical or research importance of genetic variation. As an example, Variant Effect Predictor (VEP, [6]) is applied commonly in sequencing pipelines, and uses sequence ontology (SO) terms [7] to describe a genomic variant by type of sequence alteration, genomic features altered by the change, and the predicted impact of the alteration. Pipelines may also include cross-reference to identify the variant (or variant position) in internal or external datasets that provide information used for curation of disease gene variants against recognized variant classification criteria. For example: variant frequency in outbred population groups such as gnomAD (http://gnomad.broadinstitute.org/)[8]; presence and pathogenicity assertion in ClinVar (https://www.ncbi.nlm.nih.gov/clinvar/) [9], presence in the COSMIC repository of variation identified in tumors (Catalogue of Somatic Mutations in Cancer; https://cancer.sanger.ac.uk/cosmic)[10].

References

1. Richards S, Aziz N, Bale S, et al. 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**(5):405-24 doi: 10.1038/gim.2015.30[published Online First: Epub Date]|.
2. Project) HHV. HVP Guideline HVP/GL/003-01/EN, 2016.
3. Griffith M, Spies NC, Krysiak K, et al. CIViC is a community knowledgebase for expert crowdsourcing the clinical interpretation of variants in cancer. Nat Genet 2017;**49**(2):170-74 doi: 10.1038/ng.3774[published Online First: Epub Date]|.
4. Easton DF, Deffenbaugh AM, Pruss D, et al. A systematic genetic assessment of 1,433 sequence variants of unknown clinical significance in the BRCA1 and BRCA2 breast cancer-predisposition genes. Am J Hum Genet 2007;**81**(5):873-83 doi: 10.1086/521032[published Online First: Epub Date]|.
5. Drost M, Tiersma Y, Thompson BA, et al. A functional assay-based procedure to classify mismatch repair gene variants in Lynch Syndrome. Genetics in Medicine 2018;**in press**
6. McLaren W, Gil L, Hunt SE, et al. The Ensembl Variant Effect Predictor. Genome Biol 2016;**17**(1):122 doi: 10.1186/s13059-016-0974-4[published Online First: Epub Date]|.
7. Cunningham F, Moore B, Ruiz-Schultz N, et al. Improving the Sequence Ontology terminology for genomic variant annotation. J Biomed Semantics 2015;**6**:32 doi: 10.1186/s13326-015-0030-4[published Online First: Epub Date]|.

8. Lek M, Karczewski KJ, Minikel EV, et al. Analysis of protein-coding genetic variation in 60,706 humans. Nature 2016;**536**(7616):285-91 doi: 10.1038/nature19057[published Online First: Epub Date]|.
9. Landrum MJ, Lee JM, Benson M, et al. ClinVar: improving access to variant interpretations and supporting evidence. Nucleic Acids Res 2018;**46**(D1):D1062-D67 doi: 10.1093/nar/gkx1153[published Online First: Epub Date]|.
10. Forbes SA, Beare D, Boutselakis H, et al. COSMIC: somatic cancer genetics at high-resolution. Nucleic Acids Res 2017;**45**(D1):D777-D83 doi: 10.1093/nar/gkw1121[published Online First: Epub Date]|.