ACGS Best Practice Guidelines for Variant Classification 2017

Sian Ellard[1,2], Emma L Baple[2,3,4], Martina Owens[1], Stuart Cannon[1], Diana M Eccles[5], Stephen Abbs[6], Richard Scott[4,7], Zandra C Deans[8], Tracy Lester[9], Jo Campbell[10], William G Newman[11,12], and Dominic J McMullan[13]

1. Department of Molecular Genetics, Royal Devon & Exeter NHS Foundation Trust, Exeter, EX2 5DW, UK.
2. University of Exeter Medical School, Exeter, EX2 5DW, UK.
3. Department of Clinical Genetics, Royal Devon & Exeter NHS Foundation Trust, Exeter, EX2 5DW, UK.
4. Genomics England, William Harvey Research Institute, Queen Mary University of London, Charterhouse Square, London, EC1M 6BQ, UK
5. Wessex Clinical Genetics Service, University Hospital Southampton, Southampton SO16 5YA, UK.
6. East Anglian Medical Genetics Service, Addenbrooke's Hospital, Cambridge CB2 0QQ, UK.
7. Department of Clinical Genetics, Great Ormond Street Hospital for Children NHS Foundation Trust, London, WC1N 3JH, UK.
8. UK NEQAS for Molecular Genetics, Department of Laboratory Medicine, Royal infirmary of Edinburgh, Edinburgh EH16 4SA, UK.
9. Oxford Genetic Laboratories, Oxford University Hospitals NHS Foundation Trust, Oxford OX3 7LE, UK.
10. Viapath Genetics Laboratory, Viapath Analytics LLP, 5th Floor Tower Wing, Guy's Hospital, London SE1 9RT, UK.
11. Manchester Centre for Genomic Medicine, Central Manchester University Hospitals NHS Foundation Trust, Manchester M13 9WL, UK.
12. Evolution and Genomic Science, University of Manchester, Manchester M13 9PL
13. West Midlands Regional Genetics Laboratory, Birmingham Women's NHS Foundation Trust, Birmingham, B15 2TG, UK.

Recommendations ratified by ACGS Quality Subcommittee on 2nd October 2017

1. Introduction

A rare disease is defined as a condition affecting less than one in 2000 individuals. Approximately seven thousand rare diseases have been described which in total affect an estimated 1 in 17 of the UK population (approximately 3.5 million individuals). Nearly 5000 of these rare diseases are monogenic disorders caused by highly penetrant variants in a single gene. A molecular genetic diagnosis of a rare disease requires the identification of a single disease-causing variant (or bi-allelic variants in autosomal recessive conditions). A prompt and accurate molecular diagnosis can be crucial to the delivery of optimal care for a patient and their family particularly increasingly in targeting treatment. However, diagnosis of a rare genetic disease can be a challenge and is contingent upon a robust understanding of the molecular aetiology of the disease. A molecular genetic diagnosis underpins robust disease classification, provision of prognostic information, accurate risk prediction for relatives, and importantly can indicate the most appropriate treatment(s), inform access to clinical screening, prevention strategies or clinical trials and facilitate access to support services and patient-led support groups.

Historically, genetic testing focused on the analysis of one or a small number of genes indicated by the patient's phenotype, but the advent of next generation sequencing technology has revolutionised the scale at which genetic testing can be performed enabling the analysis of many more genes within the same assay. Large gene panel tests (>100 genes) and whole exome sequencing are routinely available in UK clinical diagnostic laboratories and whole genome sequencing, currently available through the 100,000 Genomes Project in England, will be commissioned for mainstream clinical care within the NHS in the near future. Deciphering which variants are disease-causing is challenging as

each human genome has 3-4 million variants (compared to the reference human genome sequence). Only a minority are causative of monogenic disease; most are part of normal human variation or may contribute to an increased or decreased risk of multi-factorial disease. The gnomAD database (http://gnomad.broadinstitute.org/) currently includes nearly 18 million variants identified by exome sequencing of 126,216 individuals and 254 million variants identified through genome sequencing of 15,136 individuals who were part of various disease-specific and population genetic studies, (ASHG, 2016), but we do not yet have a comprehensive catalogue of global genetic variation. The focus of these guidelines is the classification of highly penetrant protein-coding variants. Inferring pathogenicity of non-coding variants is more complex, but will need to be addressed as a standard of practice in the future.

In 2015 the American College of Medical Genetics and Genomics (ACMG) and the Association for Molecular Pathology (AMP) published standards and guidelines for the interpretation of sequence variants (Richards *et al* Genetics in Medicine 2015). These guidelines describe a framework for classifying variants as "pathogenic", "likely pathogenic", "uncertain significance", "likely benign" or "benign" according to a series of criteria with levels of evidence defined as very strong, strong, moderate or supporting. They recommend that all assertions should be classified with respect to a disease and inheritance pattern. The authors note that more focused guidance regarding the classification of variants in specific genes is required given that the applicability and weight assigned to certain criteria may vary by disease and gene. Further development of these guidelines is being undertaken through the US ClinGen project which includes disease-specific working groups and a Sequence Variant Interpretation (SVI) Working Group that provides recommendations and harmonizing approaches to adapting the guidelines (see ClinGen variant pathogenicity curation webpage). The ACMG/AMP guidelines state that a variant of uncertain significance should not be used in clinical decision making. The consequences of a mis-diagnosis can be harmful not just for the proband but also their relatives whose clinical management is altered as a consequence of cascade testing.

High quality, accurate variant interpretation requires scientific knowledge of the gene structure, function, previously identified variants and disease mechanism in addition to comprehensive clinical knowledge of the patient and their families' medical history. The UK's NHS provides a unique opportunity to integrate curated genotype and phenotype information within a nationally developed database. On 4th November 2016 a group of NHS clinical scientists and clinical geneticists met to discuss the implementation of the ACMG guidelines within the UK (see workshop report from the PHG Foundation).

A consensus statement was issued on 11th November 2016 by the Association for Clinical Genomic Science (see ACGS consensus statement or Appendix 1). It recommended adoption of the ACMG guidelines for germline variant classification and interpretation in UK diagnostic genetic laboratories performing testing for rare disease and familial cancers.

A workshop entitled "Train the Trainers" was held on 28th February, 2017 and attended by representatives from 24 regional genetics centres across England, Scotland, Wales and Ireland. The overarching aim of the workshop was to plan the implementation of the ACMG guidelines in a co-ordinated way in order to achieve accurate usage and hence consistency of variant classification across and within laboratories. Worked case examples provided by

each centre were reviewed to identify areas where greater understanding on interpretation of the guidelines was required.

Please note that these guidelines are intended for general use in classifying variants in patients with rare diseases. Disease-specific guidelines are being developed for disorders where different evidence thresholds are required, for example familial breast or colorectal cancer.

2. Integration of clinical and scientific data in variant classification

Interpretation of a variant for use in clinical decision making requires comprehensive knowledge of the patient's phenotype, mode of inheritance for the disease gene, mutational mechanism (e.g. haploinsufficiency, dominant negative), protein structure/function and the strength of the gene-disease relationship (Strande *et al* 2017 bioRxiv). With the exception of the patient's phenotype data, most of this information can be obtained from the published literature/databases by a clinical scientist who can also collate the required population data and *in silico* predictions of variant effect.

The level of detailed phenotype data provided with the laboratory referral depends upon the testing scenario. If the test is for a single gene or small panel of genes where the patient's phenotype is highly specific for that disorder, then detailed phenotype data is not required for variant interpretation. In other settings, for example specialist services for heterogeneous disorders, laboratories require the submission of a defined phenotypic data set for use by clinical scientists to decide appropriate testing strategies and/or for variant interpretation. When testing is undertaken at an exome or genome scale for the diagnosis of very rare disorders, a multi-disciplinary approach is optimal, involving the referring clinician, clinical scientist and other healthcare professionals as appropriate. The purpose of the genomic multidisciplinary team (MDT) meeting is to assess the gene variant(s) identified in the context of the patient's phenotype data and ascertain their contribution to the clinical presentation.

The key question for the referring clinical team in an MDT discussion is "Does this patient's phenotype fit this gene-disease association?" If so, what is the strength of the evidence to support the variant classification? Tools to evaluate this aspect of the variant classification process are in development, for example the Summative Assessment tool within DECIPHER (https://decipher.sanger.ac.uk/). For variants of uncertain significance, the clinical team may suggest further tests that result in re-classification of the variant as "likely pathogenic" or "likely benign". These might include further genetic or non-genetic tests, clinical investigations and/or co-segregation testing.

The use of the *de novo* variant evidence criteria, PS2 and PM6, is also dependent upon phenotype assessment. These criteria **should only be used** if the patient's phenotype is consistent with the disease gene association, the nature of the testing strategy should also be considered when applying PS2 and PM6 (see Table 1 for examples of how to apply these evidence criteria with consideration given to the testing strategy employed). It is also

important to consider the possibility that variants in more than one gene are contributing to the patient's clinical presentation (Posey *et al* 2017 NEJM).

Table 1: Examples of the use of *de novo* evidence according to the type of test undertaken and the specificity of the phenotype. Note that trio exome or genome sequencing **would** reveal non-biological parental relationships.

Type of test	Parental relationships confirmed by test	Gene	Phenotype	Evidence criterion
Single gene followed by parental testing of variant	No	*NIPBL*	Classical clinical presentation of Cornelia de Lange including: Facial features, severe global developmental delay, hirsutism, upper-limb defects, failure to thrive and microcephaly	PM6
Trio exome or genome with virtual panel analysis (e.g. DDG2P in DDD study or tiered variants in 100,000 Genomes Project)	Yes	*NIPBL*	Classical clinical presentation of Cornelia de Lange including: Facial features, severe global developmental delay, hirsutism, upper-limb defects, failure to thrive and microcephaly	PS2
Gene-agnostic trio exome or genome (variants filtered by mode of inheritance)	Yes	*NIPBL*	Classical clinical presentation of Cornelia de Lange including: Facial features, severe global developmental delay, hirsutism, upper-limb defects, failure to thrive and microcephaly	PS2
Trio exome or genome with virtual panel analysis (e.g. DDG2P in DDD study or tiered variants in 100,000 Genomes Project)	Yes	*NIPBL*	Severe developmental delay; no other features of Cornelia de Lange	NOT USED
Gene-agnostic trio exome or genome (variants filtered by mode of inheritance)	Yes	*NIPBL*	Severe developmental delay; no other features of Cornelia de Lange	NOT USED
Gene panel followed by parental testing of variant	No	Many examples	Early infantile epileptic encephalopathy	PM6
Trio exome or genome with virtual panel analysis (e.g. DDG2P in DDD study or tiered variants in 100,000 Genomes Project)	Yes	Many examples	Early infantile epileptic encephalopathy	PS2_Moderate
Gene-agnostic trio exome or genome (variants filtered by mode of inheritance)	Yes	Many examples	Early infantile epileptic encephalopathy	PS2_Moderate
Trio exome or genome with virtual panel analysis (e.g. DDG2P in DDD study or tiered variants in 100,000 Genomes Project)	Yes	Many examples	Non-syndromic Intellectual disability	PS2_Supporting
Gene-agnostic trio exome or genome (variants filtered by mode of inheritance)	Yes	Many examples	Non-syndromic Intellectual disability	PS2_Supporting

PP4 can be used as a supporting piece of evidence when the patient's phenotype in its entirety is consistent with a specific single genetic aetiology. The third evidence criterion that may utilise phenotypic information is PS3. This can be used as strong evidence if there is a specific aspect of the phenotype that is measurable using biochemical, mRNA studies or other clinical investigations and the results are pathognomonic of a specific genetic cause of a disorder (see Table 2 below for examples).

Table 2: Examples of functional tests that could be used as evidence for PS3 (Well-established *in vitro* or *in vivo* functional studies supportive of a damaging effect on the gene or gene product).

Type of test	Example
Biochemical	Enzyme level within range that is diagnostic for the disease e.g. low/absent alpha-galactosidase levels in Fabry disease
Muscle biopsy	Evidence of a specific pathology, defined by immunohistochemistry or direct morphological assessment, associated with a single genetic aetiology
Drug response	Improved glycaemic response in patients with HNF1A/4A MODY treated with sulphonylurea tablets
Messenger RNA analysis	Aberrant splicing shown by reverse transcriptase-PCR and characterisation of product(s) by Sanger sequencing
Microsatellite analysis of tumour tissue	Somatic loss of heterozygosity at the *MEN1* locus in a patient with a germline novel variant
Neuroimaging	Findings consistent with a single specific genetic aetiology (e.g. a *CASK* mutation)
Haematological	HbH inclusion bodies in ATRX syndrome
Nerve conduction studies	Prolongation of distal nerve conduction latencies in an individual with clinical features consistent with hereditary neuropathy with liability to pressure palsies

Examples of test results that should not be used as strong evidence are given in Table 3.

Table 3: Examples of tests that should not be used as evidence for PS3 (Well-established *in vitro* or *in vivo* functional studies supportive of a damaging effect on the gene or gene product).

Type of test	Example
Biochemical test	Raised serum or CSF lactate indicative of mitochondrial disorder but not specific to any one particular disorder
Neuroimaging	Absent/hypoplastic corpus callosum is seen in a number of genetic disorders and therefore is not a specific finding
Nerve conduction studies	Nerve conduction velocities supporting a diagnosis of Hereditary motor and sensory neuropathy but not specific for a particular subtype
Muscle biopsy	Non-specific myopathic changes

In certain circumstances, for example where a hypothesis-free whole exome or genome testing strategy has been used, it may be appropriate to use the results of a measurable test such as nerve conduction studies or neuroimaging findings that would be specific to a limited number of genetic disorders as strong evidence using PS3. This may be deemed reasonable because the test strategy does not target a specific phenotype driven gene selection and because the other genetic causes of the measurable test result should have been excluded

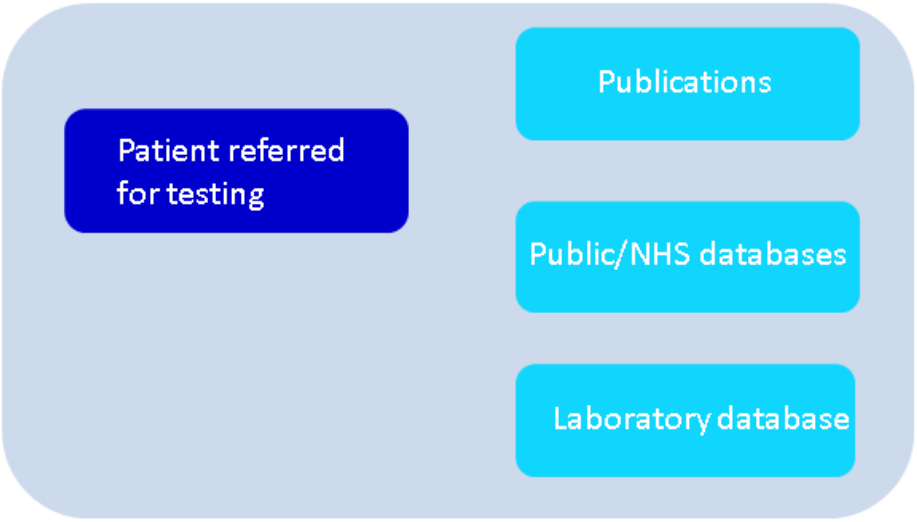
as part of the genomic testing. All such decisions should only be made after careful multidisciplinary discussion.

The ACMG/AMP variant classification guidelines may also be applied in interpreting sequence data from patients with common disease phenotypes where the purpose is to identify high penetrance genetic predisposition. Examples include familial breast or colorectal cancer, inherited cardiac conditions and monogenic diabetes. Phenotype and/or family history data are used to estimate the prior probability of a single highly penetrant gene accounting for the majority of the phenotype. Phenotypic information is often used to select patients for genetic testing but additional information to underpin a robust interpretation will often be lacking in the absence of a family history. Caution is needed since (benign) rare variants and common phenotypes may coincide frequently, phenocopies are common and other genetic and environmental factors influence penetrance and phenotype in gene carriers and non-carriers. As noted above, different evidence thresholds may be required in these disorders and disease-specific guidelines are being developed for familial cancers and cardiac disorders. We note that where lower penetrance genes or genetic variants are included in a gene panel test, any lower penetrance pathogenic variant(s) identified are unlikely to account for the majority of the phenotype/risk and this should be clearly articulated.

3. Variant classification: Supplementary notes for use of the ACMG evidence criteria

The assessment of a variant should include phenotype data from all patients currently identified with the variant; the patient referred for testing, previous patients tested in the laboratory, published literature and information from variant databases (see Figure 1).

Figure 1: The evidence for a variant classification is assessed across all patients for which information is available.



The framework developed by the ACMG team utilises a series of evidence criteria in support of a pathogenic (P) or benign (B) classification which are described in tables 3 and 4 of the publication by Richards *et al* (2015). In Table 4 (below) we describe additional information to assist with the application of the ACMG guidelines. These notes must be used in conjunction with the detailed guidance published by Richards *et al* (2015) and Jarvik & Browning (2016). The principles of Bayes' theorem apply to variant classification in that each item of evidence in support of or against pathogenicity should be used only once.

Table 4: Supplementary information for classifying pathogenic (P) or benign (B) variants

Evidence criteria (level) supplementary notes
PVS1 – (Very Strong) null variant (nonsense, frameshift, canonical ±1 or 2 splice sites, initiation codon, single or multi-exon deletion) in a gene where LOF is a known mechanism of disease. Note that caution is required when interpreting 3' nonsense or frameshift variants predicted to escape nonsense mediated decay and consensus spice donor/acceptor site variants predicted to lead to in frame deletions. For example the *BRCA2* nonsense variant, p.(Lys3326Ter) c.9976A>T, results in loss of the last 93 amino acids of the BRCA2 protein but does not confer a high risk of familial breast cancer (Mazoyer *et al* 1996). Nor does the *BRCA1* c.594-2A>C slice acceptor site variant (de la Hoya *et al* 2016).
PS1 – (Strong) Same amino acid change as a previously established pathogenic variant regardless of nucleotide change. This criterion can be used if there is sufficient evidence for pathogenicity for the same missense variant (ie an amino acid change) caused by a **different** base substitution. For example the previously reported variants is p.Val12Leu (c.34G>**C**) and your patient's variant is p.Val12Leu (c.34G>**T**) as described by Richards *et al* (2015).
PS2 – (Strong) De novo (both maternity and paternity confirmed) in a patient with the disease and no family history. Note that the genotype **must be consistent** with the phenotype. If a *de novo* variant was identified by trio exome or genome sequencing then maternity and paternity **already** should have both been confirmed by using a bioinformatics pipeline that would reveal inconsistencies with inheritance. In the situation that a *de novo* variant is identified by trio exome or genome sequencing a cautious approach is recommended (since every exome typically contains between 1-2 *de novo* non-synonymous variant and the testing strategy that has been employed will identify these). If the patient's phenotype is non-specific or there is evidence of significant genetic heterogeneity (e.g. intellectual disability), this criterion should only be used at a lower level. Please see Table 1 for examples.
PS3 – (Strong) Well-established in vitro or in vivo functional studies supportive of a damaging effect on the gene or gene product. Functional studies can include *in vitro* functional assays (eg. reporter gene assays for transcription factors), measurement of proteins *in vivo* (e.g. biochemical tests on patient samples), mRNA analysis for suspected splicing variants and other investigations where the results are pathognomonic of a **specific genetic cause** of a disorder. See Table 2 for a list of examples. Note that evidence from functional studies must be carefully assessed to determine the data quality, reliability and hence degree of confidence in the results. For example a test that is carried out in a certified diagnostic laboratory, has been replicated in a second centre, or a variant that has undergone multiple functional assessments using different methodologies would provide greater confidence that the variant has a damaging effect upon the gene product.
PS4 – (Strong) The prevalence of the variant in affected individuals is significantly increased compared with the prevalence in controls Case control study data is rarely available for rare diseases, but PS4 can be used as a **moderate** level of evidence if the variant has been *previously identified in multiple (two or more) unrelated affected individuals* and has not been reported in gnomAD (see Note 2 in Table 3, Richards *et al* 2015). PS4 can be used as a **supporting** level of evidence if the variant has been *previously identified in one unrelated affected individual* and has not been reported in gnomAD. In practice this is most applicable to autosomal dominant disorders. Absence from the gnomAD database also allows use of PM2 at moderate level, i.e. both PS4 (moderate or supporting) and PM2 can be used.
PM1 – (Moderate) Located in a mutational hot spot and/or critical and well-established functional domain (e.g. active site of an enzyme) without benign variation. It can be useful to plot functional domains, gnomAD variants, HGMDPro/ClinVar/LOVD etc variants and Consurf/conservation plots to show reported pathogenic variants, proxy population and benign variants and amino acid conservation for a region of a gene (see Figures 2, 3 and this link). *In silico* protein modelling data can

be included as supporting evidence.

PM2 – (Moderate) Absent from controls (or at extremely low frequency if recessive) in Exome Aggregation Consortium.
It is important to check that the variant position is covered to sufficient read depth in ExAC (or gnomAD). To check in ExAC see this link. Be aware that indels are less readily identified by next generation sequencing and ascertain whether other indels have been detected within the region.

PM4 – (Moderate) Protein length changes as a result of in-frame deletions/insertions in a non-repeat region or stop-loss variants.
This criterion is used for in-frame deletions or insertions and would also apply to a deletion of a small in-frame exon. PVS1 is used for out of frame exon deletions and larger in-frame exon deletions that remove a significant proportion of a gene. There is no fixed definition of small/large as the impact of a deletion will depend on the size of a gene and the gene architecture (including the impact of a deletion on functional domains or regulatory elements). Greater care should be taken with apparent in-frame exonic insertions/duplications since it is harder to predict their impact at the protein level, and their precise location and orientation may not be known unless demonstrated by whole genome sequencing.

PP1 – (Supporting) Co-segregation with disease in multiple affected family members in a gene definitively known to cause the disease.
The thresholds suggested by Jarvik and Browning (2016) should be used. It is important to consider the number of meioses, **not** the number of informative individuals. Incomplete penetrance, age of onset and phenocopy rates can be incorporated within the calculation.

PP2 – (Supporting) PP2 Missense variant in a gene that has a low rate of benign missense variation and in which missense variants are a common mechanism of disease.
ExAC constraint scores can be used as evidence for a low rate of benign variation (Lek *et al* 2016). Z scores ≥3.09 (marked amber in ExAC) are significant but it is important to consider constraint for the region encompassing the variant, not just across the entire gene.

PP3 – (Supporting) Multiple lines of computational evidence support a deleterious effect on the gene or gene product (conservation, evolutionary, splicing impact, etc.).
In silico splicing prediction tools can be used as evidence to suggest a significant impact on splicing potential for splice site variants outside the canonical splice acceptor (-1 and -2) and donor (+1 and +2) regions. PS3 can be used if mRNA analysis is undertaken and demonstrates the presence of an abnormal transcript(s) predicted to result in loss of protein expression. In this situation PP3 would not apply as well.
For predicting the impact of missense variants it is likely that a meta-predictor tool (e.g. REVEL, Ioannidis *et al* 2016) will replace the use of multiple prediction tools that each assess a single line of evidence.

PP4 – (Supporting) Patient's phenotype or family history is highly specific for a disease with a single genetic aetiology.
For example a patient with a clinical diagnosis of Marfan syndrome defined by the Ghent clinical criteria (Loeys *et al* 2010) who has undergone *FBN1* testing.

This evidence criterion incorporates the prior probability that a patient will have a pathogenic variant in a particular gene or genes and therefore does not need to be limited to diseases where there is a single genetic aetiology. This criterion may be applied in the scenario where a patient has a rare combination of clinical features for which there are a very limited number of known genetic aetiologies and all those genes have been tested. **The key consideration is the specificity of the phenotype.** Caution should be taken when considering phenotypic features which are specific to a disorder that is genetically heterogeneous. Non-specific phenotypes such as intellectual disability, seizure disorder without a specific EEG pattern and subtle abnormalities of the corpus callosum should never be used in isolation as evidence for PP4.

The testing strategy used to identify the variant is also important. For example, when a single gene test has been undertaken because the patient's phenotype is a "good fit" for that specific genetic aetiology, there is a high prior probability that a variant identified within that gene will be causative of the patient's disease and the test specificity is high. In contrast, when a large panel test for a genetically heterogeneous condition is performed, the overall prior probability for finding a causative variant is the sum of the prior probabilities for each individual gene. Some genes may account for only a small proportion of cases (low prior probability of being the disease gene responsible for the patient's phenotype) but are highly polymorphic (hence a high chance of a novel/rare benign variant being found). In general the more genes tested, the higher the likelihood of finding a spurious variant not causal of the patient's phenotype. Hence the test specificity is reduced compared to a single gene test. Using a gene-agnostic whole exome or genome sequencing strategy with variant filtering by mode of inheritance provides significantly increased specificity compared to a gene panel approach and can be cited as additional evidence.

PP5 – (Supporting) Reputable source recently reports variant as pathogenic, but the evidence is not

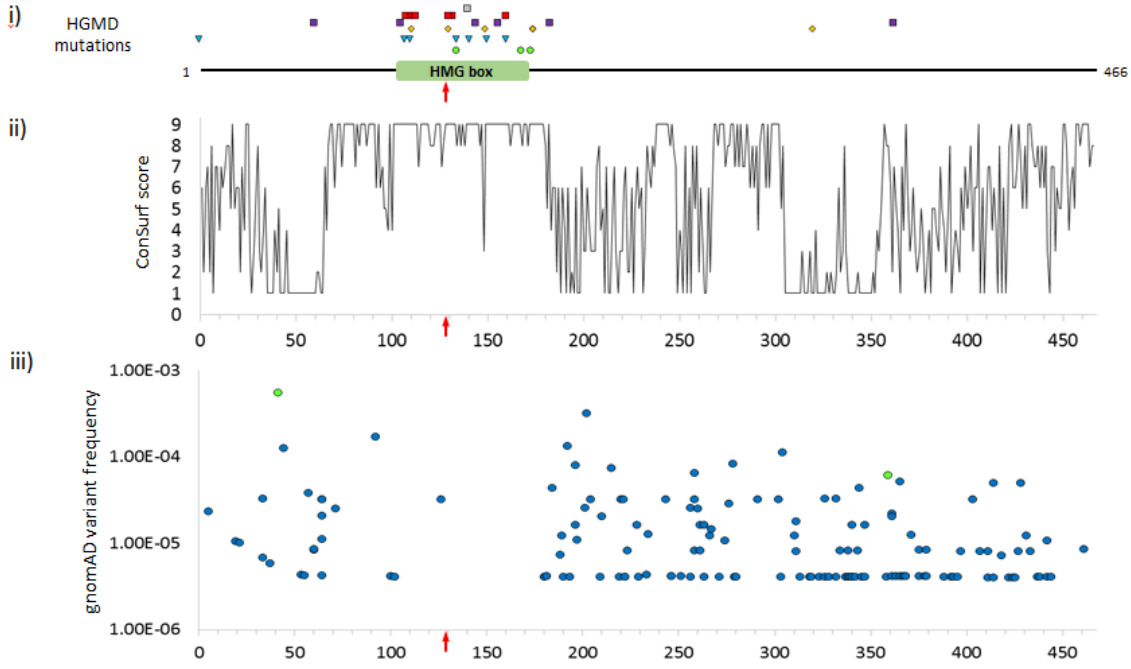
available to the laboratory to perform an independent evaluation.
This will only be used in exceptional circumstances as the expectation is to seek out the information supporting the classification. For example, if a variant has been reported on ClinVar or DECIPHER, the laboratory should contact the originating laboratory/clinician to obtain information about the inheritance and phenotype. This can be mutually beneficial as the additional evidence you share with the originating laboratory may allow them to re-classify the variant for their patient's benefit.

BS1 – (Strong) Allele frequency is greater than expected for disorder.
A very useful tool is available to determine whether the allele frequency of the variant is greater than expected for the disorder (Whiffin *et al* 2016). In the absence of precise information about the disease prevalence and penetrance we recommend using conservative settings (by selecting the highest likely prevalence and the lowest likely penetrance) to see if the variant frequency on the gnomAD database exceeds the maximum credible allele frequency. The tool can be accessed at http://cardiodb.org/allelefrequencyapp/. *For an autosomal dominant disorder with high penetrance it is acceptable to use BS1_Strong as stand-alone evidence to classify a variant as likely benign.*

BP1 – (Supporting) Missense variant in a gene for which primarily truncating variants are known to cause disease.
This criterion can also be used for loss of function variants in a gene where the disease is caused by gain of function variants or dominant negative loss of function variants (e.g. those in the last exon of a gene).

Figure 2: Example plot of *SOX10* functional domains, reported pathogenic variants, proxy population/benign variants and amino acid conservation for a region of a gene plots to support evidence criterion PM1. The red arrow indicates the position of the p.Ala132 residue (see variant classification evidence in Figure 4).

(i) The 466 residue sequence of SOX10 is shown as a black horizontal line, with the position of the HMG box (104-172, UniProtKB annotation) shown by a green box; above the line are shown positions of HGMD mutations in Waardenburg syndrome, type unspecified (grey square), Waardenburg syndrome type 2 (red squares), Waardenburg syndrome type 4 (purple squares), PCWH (yellow rhomboids), Kallmann syndrome (with or without deafness; blue triangles), or other phenotype (green circles).
(ii) ConSurf grade, from 1 (minimum) to 9 (maximum), plotted by amino acid position.
(iii) gnomAD missense variants, plotted by position (*x*-axis) against allele frequency (*y*-axis); green circles represent variants which have been observed in the homozygous state.



Courtesy of Dr Richard Caswell, University of Exeter Medical School

Figure 3: Example plot of *SOX10* functional domains, ClinVar variants, proxy population/benign variants and missense constraint from DECIPHER (https://decipher.sanger.ac.uk/gene/SOX10#overview/protein-info).

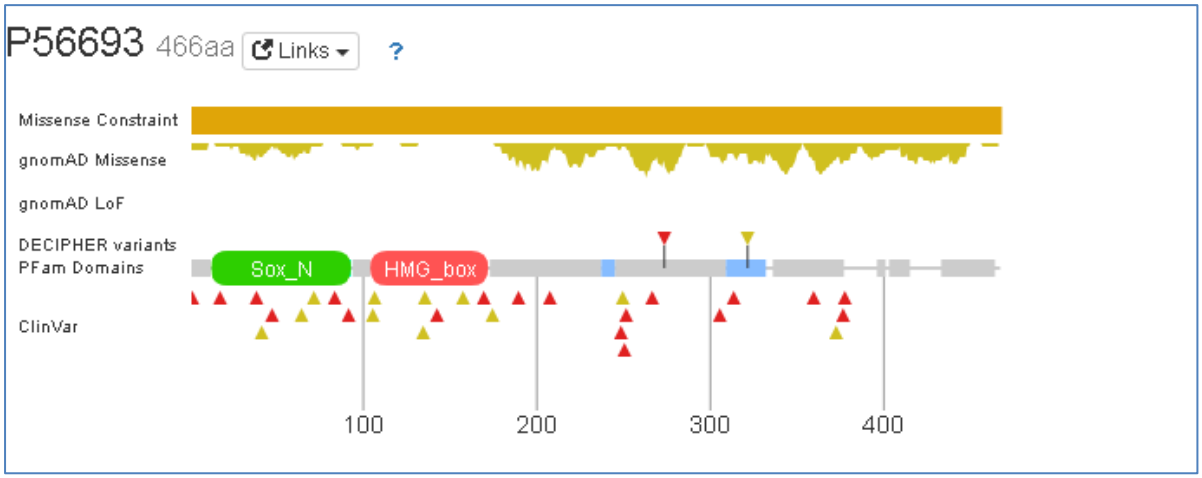


Figure 4. Evidence for *SOX10* p.(Ala132Val) classification

Gene:	***SOX10***
Variant description: HGVS	**NM_006941.3:c.395C>T p.(Ala132Val)**
Variant location: GRCh37 (hg19):	**Chr22:g.38379397**
Variant classification: (Pathogenic/Likely pathogenic/ Uncertain significance/Likely benign/Benign)	**Likely Pathogenic**

Evidence for variant classification (ACMG code – level)
• The p.Ala132 residue is located within the HMG box DNA binding domain. *In silico* comparative modelling predicts that this residue interacts with DNA and that the substitution of alanine with valine disrupts the DNA-binding surface (PM1_Moderate). ▪ The p.(Ala132Val) variant is not listed in the ExAC database (60,706 individuals) nor in gnomAD (138,632 individuals) (PM2_Moderate). ▪ The *SOX10* gene has a low rate of benign missense variation as evidenced by a significant (z= 3.31) ExAC constraint score (PP2_Supporting). ▪ The p.Ala132 residue is conserved across 20 species (to Zebrafish) and the Consurf score is 9. The p.(Ala132Val) variant is predicted by SIFT, PolyPhen and AlignGVGD to have a deleterious effect on protein function (PP3_Supporting). ▪ The *SOX10* variant was identified by trio exome analysis using a gene agnostic approach and the

patient's phenotype is consistent with Waardenburg syndrome (PP4_Supporting).

4. **Reporting the variant classification**

The aim of genomic testing for a patient with a rare disease of unknown cause is to provide a genetic diagnosis by identifying a (likely) disease-causing variant (or bi-allelic variants in autosomal recessive conditions). The genetic analysis may involve the pathogenic assessment of one or multiple variants but the clinical diagnostic report will only describe those that are relevant, or have likely or possible relevance to the patient's clinical presentation. Results included within a clinical diagnostic report will form part of the patient's clinical record and should be unambiguous to a non-specialist.

Variants are classified as "pathogenic", "likely pathogenic", "uncertain significance", "likely benign" or "benign" with respect to a disease and inheritance pattern. The variant classification must be included within the main body of the clinical diagnostic report together with the gene-disease association and the mode of inheritance. The evidence supporting the variant classification can be included in an appendix to the report (see example in Figure 4).

Next generation sequencing of large gene panels, a whole exome or whole genome will identify multiple variants of uncertain significance. For large gene panels the more genes tested, the higher the likelihood of finding spurious variants not causal of the patient's phenotype. It is important to remember that not all types of variants can be detected by next generation sequencing and there are still many gene-disease associations to be discovered. The consequences of a mis-diagnosis have wider family implications beyond the proband.

The Association for Clinical Genomic Science (ACGS) Practice guidelines for Targeted Next Generation Sequencing Analysis and Interpretation recommend that variants of uncertain significance are only reported within the main body of the report if there is further genetic testing or clinical investigation that is likely to re-classify the variant as either "likely pathogenic" or "likely benign". It is important to consider the context in which genomic testing is undertaken. For example rapid testing of cancer predisposition genes such as *BRCA1* or *BRCA2* to determine pharmaceutical treatment options for the affected patient might report only "pathogenic" or "likely pathogenic" variants, in contrast to testing that seeks to inform prophylactic surgery options for the proband and their relatives where further testing/investigation of a variant of uncertain significance may be appropriate. However, there may also be situations in which an MDT discussion concludes that there is clinical utility in reporting a variant of uncertain significance where it may be impossible to obtain sufficient evidence to reach a variant classification of "likely pathogenic" due to lack of family member samples, but all the available clinical, gene-level and variant-level evidence supports the likely diagnosis. For example, a child with a developmental disorder where no cascade or prenatal testing is required, but a likely diagnosis will provide access to support services for the patient and their family.

The ACGS recommends that unless there are specific reasons to the contrary (examples as described above), the following statement is included in a report for a variant of uncertain significance **"This result does not confirm a genetic diagnosis of *disorder X* and should not be used in isolation for clinical decision making".**

5. **Reclassification of variants**

Variant data and relevant associated information must be stored within the laboratory in a way that allows reclassification if required. Sharing of variant data on a global scale in a manner that conforms to UK information governance requirements is a goal supported by the ACGS and BSGM (British Society for Genetic Medicine). The DECIPHER database (https://decipher.sanger.ac.uk/) hosts an NHS consortium project to allow sharing of variant data in a restricted manner. This allows member laboratories to identify conflicting classifications for the same variant to enable submitters to discuss the most appropriate classification based on the available evidence.

Reassessment of a variant that results in reclassification may be prompted by the publication of new knowledge regarding the variant (or gene-disease association); by a request for a family member test or as a result of further clinical investigations or evolution of the patient's phenotype that questions the original diagnosis.

We propose that reclassification of a variant across categories that fundamentally changes the clinical relevance – i.e. not from likely benign to benign or likely pathogenic to pathogenic (or vice versa) – should be shared with other relevant health care professionals. The laboratory where the new information is generated should liaise with any laboratories (where this is feasible) that generated the original classification status to ensure consistency across centres. The new classification data and the basis for this classification should be placed in a publicly accessible database so that the information is available widely. If the new classification has potential importance for clinical management e.g. classification of a *BRCA1* or *BRCA2* variant that may alter decisions around risk reducing mastectomy, this decision should be documented (through an MDT) and communicated to the patient's clinical team(s) as quickly as possible. If such a variant is identified, in addition to the above two actions, this variant classification information should be disseminated rapidly (within one week) to all other diagnostic genetics laboratories within the UK using a designated secure NHS e-mail address for each laboratory. It is hoped that in the near future a more regulated method for data-sharing will be created for NHS Genomic laboratories, but ultimately it is the professional responsibility of the Clinical Genomic community to ensure data is shared responsibly for improved patient care.

Acknowledgements

We are grateful to Dr Helen Firth (Addenbrooke's Hospital, Cambridge, UK), Dr Heidi Rehm and Dr Steven Harrison (Harvard, USA) for extremely helpful discussions and all their support.

References

de la Hoya M, Soukarieh O, López-Perolio I, Vega A, Walker LC, van Ierland Y, Baralle D, Santamariña M, Lattimore V, Wijnen J, Whiley P, Blanco A, Raponi M, Hauke J, Wappenschmidt B, Becker A, Hansen TV, Behar R, Investigators K, Niederacher D, Arnold N, Dworniczak B, Steinemann D, Faust U, Rubinstein W, Hulick PJ,

Houdayer C, Caputo SM, Castera L, Pesaran T, Chao E, Brewer C, Southey MC, van Asperen CJ, Singer CF, Sullivan J, Poplawski N, Mai P, Peto J, Johnson N, Burwinkel B, Surowy H, Bojesen SE, Flyger H, Lindblom A, Margolin S, Chang-Claude J, Rudolph A, Radice P, Galastri L, Olson JE, Hallberg E, Giles GG, Milne RL, Andrulis IL, Glendon G, Hall P, Czene K, Blows F, Shah M, Wang Q, Dennis J, Michailidou K, McGuffog L, Bolla MK, Antoniou AC, Easton DF, Couch FJ, Tavtigian S, Vreeswijk MP, Parsons M, Meeks HD, Martins A, Goldgar DE, Spurdle AB. Combined genetic and splicing analysis of BRCA1 c.[594-2A>C; 641A>G] highlights the relevance of naturally occurring in-frame transcripts for developing disease gene variant classification algorithms. Hum Mol Genet. 2016 Jun 1;25(11):2256-2268. PubMed PMID: 27008870.

Jarvik GP, Browning BL. Consideration of Cosegregation in the Pathogenicity Classification of Genomic Variants. Am J Hum Genet. 2016 Jun 2;98(6):1077-81. doi: 10.1016/j.ajhg.2016.04.003. Epub 2016 May 26. PubMed PMID: 27236918.

Ioannidis NM, Rothstein JH, Pejaver V, Middha S, McDonnell SK, Baheti S, Musolf A, Li Q, Holzinger E, Karyadi D, Cannon-Albright LA, Teerlink CC, Stanford JL, Isaacs WB, Xu J, Cooney KA, Lange EM, Schleutker J, Carpten JD, Powell IJ, Cussenot O, Cancel-Tassin G, Giles GG, MacInnis RJ, Maier C, Hsieh CL, Wiklund F, Catalona WJ, Foulkes WD, Mandal D, Eeles RA, Kote-Jarai Z, Bustamante CD, Schaid DJ, Hastie T, Ostrander EA, Bailey-Wilson JE, Radivojac P, Thibodeau SN, Whittemore AS, Sieh W. REVEL: An Ensemble Method for Predicting the Pathogenicity of Rare Missense Variants. Am J Hum Genet. 2016 Oct 6;99(4):877-885. PubMed PMID.

Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, O'Donnell-Luria AH, Ware JS, Hill AJ, Cummings BB, Tukiainen T, Birnbaum DP, Kosmicki JA, Duncan LE, Estrada K, Zhao F, Zou J, Pierce-Hoffman E, Berghout J, Cooper DN, Deflaux N, DePristo M, Do R, Flannick J, Fromer M, Gauthier L, Goldstein J, Gupta N, Howrigan D, Kiezun A, Kurki MI, Moonshine AL, Natarajan P, Orozco L, Peloso GM, Poplin R, Rivas MA, Ruano-Rubio V, Rose SA, Ruderfer DM, Shakir K, Stenson PD, Stevens C, Thomas BP, Tiao G, Tusie-Luna MT, Weisburd B, Won HH, Yu D, Altshuler DM, Ardissino D, Boehnke M, Danesh J, Donnelly S, Elosua R, Florez JC, Gabriel SB, Getz G, Glatt SJ, Hultman CM, Kathiresan S, Laakso M, McCarroll S, McCarthy MI, McGovern D, McPherson R, Neale BM, Palotie A, Purcell SM, Saleheen D, Scharf JM, Sklar P, Sullivan PF, Tuomilehto J, Tsuang MT, Watkins HC, Wilson JG, Daly MJ, MacArthur DG; Exome Aggregation Consortium.. Analysis of protein-coding genetic variation in 60,706 humans. Nature. 2016 Aug 18;536(7616):285-91. doi: 10.1038/nature19057. PubMed PMID: 27535533.

Loeys BL, Dietz HC, Braverman AC, Callewaert BL, De Backer J, Devereux RB, Hilhorst-Hofstee Y, Jondeau G, Faivre L, Milewicz DM, Pyeritz RE, Sponseller PD, Wordsworth P, De Paepe AM. The revised Ghent nosology for the Marfan syndrome. J Med Genet. 2010 Jul;47(7):476-85. doi: 10.1136/jmg.2009.072785. PubMed PMID:20591885.

Mazoyer S, Dunning AM, Serova O, Dearden J, Puget N, Healey CS, Gayther SA, Mangion J, Stratton MR, Lynch HT, Goldgar DE, Ponder BA, Lenoir GM. A polymorphic stop codon in BRCA2. Nat Genet. 1996 Nov;14(3):253-4. DOI: 10.1038/ng1196-253 PMID: 8896551

Posey JE, Harel T, Liu P, Rosenfeld JA, James RA, Coban Akdemir ZH, Walkiewicz M, Bi W, Xiao R, Ding Y, Xia F, Beaudet AL, Muzny DM, Gibbs RA, Boerwinkle E, Eng CM, Sutton VR, Shaw CA, Plon SE, Yang Y, Lupski JR. Resolution of Disease Phenotypes Resulting from Multilocus Genomic Variation. N Engl J Med. 2017 Jan 5;376(1):21-31. doi: 10.1056/NEJMoa1516767. Epub 2016 Dec 7. PubMed PMID: 27959697.

Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, Grody WW, Hegde M, Lyon E, Spector E, Voelkerding K, Rehm HL; ACMG Laboratory Quality Assurance Committee. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med. 2015 May;17(5):405-24. doi: 10.1038/gim.2015.30. Epub 2015 Mar 5. PubMed PMID: 25741868.

Strande NT, Rooney Riggs E, Buchanan AH, Ceyhan-Birsoy O, Dwight ST, Goldstein JL, Ghosh R, Seifert BA, Sneddon TP, Wright MW, Milko LV, Giovanni MA, Murray MF, O'Daniel JM, Ramos EM, Santani AB, Scott AF, Cherry JM, Plon SE, Rehm HL, Lese Martin C, Berg JS; Evaluating the clinical validity of gene-disease associations: an evidence-based framework developed by the Clinical Genome Resource. bioRxiv 2017 https://doi.org/10.1101/111039.

Whiffin N, Minikel E, Walsh R, O'Donnell-Luria A, Karczewski K, Ing AY, Barton PJR, Funke B, Cook SA, MacArthur DG, Ware JS; Using high-resolution variant frequencies to empower clinical genome interpretation. bioRxiv 2016 https://doi.org/10.1101/073114.

Appendix 1: ACGS Consensus Statement 11/11/2016

Consensus statement on adoption of American College of Medical Genetics and Genomics (ACMG) guidelines for sequence variant classification and interpretation 11/11/2016

Headline Consensus Statement

ACGS recommends adoption of the ACMG guidelines (Richards, 2015) for sequence variant classification and interpretation in UK diagnostic genetic laboratories carrying out testing for rare disease and familial cancers.

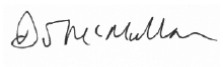
Background

Classification and interpretation of genomic variation is a highly complex discipline and in the clinical setting the need for accuracy and consistency is essential to maximise patient benefit and minimise harm. The revolution in genomic technology has led to increased routine detection of novel variants in a rapidly increasing number of novel disease genes. ACMG recently attempted to address the challenges faced by devising a detailed systematic framework for sequence variant interpretation which has now been widely adopted in the US and many European centres. Furthermore, expert panels are being formed as part of the ClinGen resource consortium to develop gene and disease specific criteria to supplement the original framework.

A Workshop hosted by ACGS was held at Austin Court, Birmingham on 4th November 2016 to reach an expert consensus view on adoption of the ACMG guidelines by the UK clinical genomics community. 70 delegates attended representing most Regional Genetics services (lab and clinical teams) with additional invited representation from all BSGM constituent groups, NHS-E (Genomics Implementation Unit), Genomics England, UKNEQAS, UKGTN, DECIPHER/DDD, HEE (Genomics Education Programme) and PHG-Foundation. The agenda included presentations on NEQAS assessment of consistency in interpretation, experiences in early adoption of ACMG guidelines, harmonisation with CNV classification, frameworks for clinical/phenotypic classification, and integration of ACMG framework into DECIPHER, together with breakout group discussions. A summary report of the Workshop is being prepared by PHG-Foundation for circulation in early 2017.

There was clear consensus agreement that the UK clinical genomics community should adopt ACMG sequence interpretation guidelines as soon as possible. ACGS, with support from HEE, will develop a multi-disciplinary training approach starting with a train-the-trainer event in early 2017. This will catalyse centre and region based adoption and also identify UK specific issues which may augment the guidelines when they are built into UK Best Practice guidelines. ACGS and BSGM annual meetings in 2017 will incorporate updates on progress with these important initiatives. In addition ACGS will offer to support further guideline development with ClinGen and ACMG via distributed variant data analyses.

Signed by: **Dom McMullan | Chair, ACGS**



Endorsed by: **Professor Bill Newman | Chair, BSGM**