

Tailoring the ACMG and AMP guidelines for the interpretation of sequenced variants in the *FBN1* gene for Marfan syndrome: Proposal for a disease- and gene-specific guideline.

1st author and running title: Muiño-Mosquera. *FBN1*-specific variant interpretation guidelines

List of co-authors:

Laura Muiño-Mosquera, MD ^{1,2*}

Felke Steijns, MSc ¹

Tjorven Audenaert, MD ³

Ilse Meerschaut, MD ¹

Anne De Paepe, MD, PhD¹

Wouter Steyaert, BSc ¹

Sofie Symoens, MSc, PhD ¹

Paul Coucke, MSc, PhD¹

Bert Callewaert, MD, PhD¹

Marjolijn Renard, MSc, PhD^{1#}

Julie De Backer, MD, PhD ^{1,3 #}

¹ Center for Medical Genetics, Ghent University Hospital, Ghent, Belgium

² Division of Pediatric Cardiology, Ghent University Hospital, Ghent, Belgium

³ Department of Cardiology, Ghent University Hospital, Ghent, Belgium

Joint last co-authors

*Corresponding author

Dr. Laura Muiño-Mosquera, Center for Medical Genetics, Ghent University Hospital

De Pintelaan 185, 9000, Belgium.

Tel: (+32) (0) 9 332 2441.

Fax: (+32) (0) 9 332 4970

Email: laura.muinomosquera@uzgent.be

Abstract

Background: The introduction of next-generation sequencing techniques has substantially increased the identification of new genetic variants and hence the necessity of accurate variant interpretation. In 2015 the American College of Medical Genetics and Genomics and the Association for Molecular Pathology (ACMG/AMP) proposed new variant interpretation guidelines. Gene specific characteristics were, however, not considered, sometimes leading to inconsistent variant interpretation.

Methods: To allow a more uniform interpretation of variants in the *FBN1* gene, causing Marfan syndrome, we tailored these guidelines to this gene and disease. We adapted 15 of the 28 general criteria and classified 713 *FBN1* variants previously identified in our laboratory as causal mutation or variant of uncertain significance (VUS) according to these adapted guidelines. We then compared the agreement between previous methods and the adapted ACMG/AMP criteria.

Results: Agreement between the methods was 86.4% (K-alpha 0.6). Application of the tailored guidelines resulted in an increased number of VUSs (14.5% to 24.2%). Of the 85 variants that were downscaled to likely benign or VUS, 59.7% were missense variants outside a well-established functional site. Available clinical- or segregation data, necessary to further classify these types of variants, were in many cases insufficient to aid the classification.

Conclusion: Our study shows that classification of variants remains challenging and may change over time. Currently, a higher level of evidence is necessary to classify a variant as pathogenic. Gene-specific guidelines may be useful to allow a more precise and uniform interpretation of the variants in order to accurately support clinical decision-making.

Keywords

Marfan syndrome, Fibrillin-1, ACMG/AMP guidelines, variant classification, variant interpretation.

INTRODUCTION

Next generation DNA sequencing has markedly advanced mass data generation at lower cost. As a consequence of easily accessible large gene panel testing or exome/genome analysis, considerable numbers of variants that need correct interpretation are identified. The latter is an arduous process, especially when clinical data required to correlate a genotype with the phenotype are lacking. The significance of identified variants can range from certainly pathogenic to certainly benign, but in many cases the clinical significance of these variants remains uncertain. Furthermore, variant interpretation can differ between laboratories(1). For clinical purposes, precise classification and interpretation is, however, essential in the diagnostic, therapeutic and genetic counseling processes.

In 2015 the American College of Medical Genetics and Genomics and the Association for Molecular Pathology (ACMG/AMP) released new general guidelines to aid the correct classification and interpretation of variants in Mendelian diseases(2). First, the guidelines recommend indicating the level of pathogenicity, as “pathogenic” (P), “likely pathogenic” (LP), “variant of uncertain significance” (VUS), “likely benign” (LB) or “benign” (B). The terms LP and LB indicate that a certain variant is 90% likely to be pathogenic or benign, respectively. Second, the guidelines provide a methodology to allow for this variant classification. This methodology includes an evaluation of seven different categories: 1. the prevalence of the variant in population databases, 2. genotype-phenotype data, 3. literature and locus specific databases, 4. computational and predictive data, 5. functional data, 6. co-segregation analysis and 7. allelic data. Several levels of evidence were defined for each of these categories, i.e. supportive (P), moderate (M), severe (S) and very severe (VS) resulting in a total of 28 different criteria which can be evaluated. Subsequently these criteria can be combined according to a set of rules to reach a decision on the level of pathogenicity(2)

Although the ACMG/AMP guidelines represent a major step forward to the standardization of variant classification, the practical application of these guidelines remains challenging because some of the proposed criteria are still open to subjective interpretation(3–5). To outbalance this problem several initiatives have been undertaken: (1)1. Different computational programs have been developed to facilitate the use of the ACMG/AMP guidelines. These tools can reduce human errors, enable critical evaluation of pathogenicity and detect and resolve discordant conclusions(5–7). Importantly, these tools lack gene-specific knowledge. (2)2. Another initiative has been the refinement of the ACMG/AMP guidelines by using 108 detailed specifications and a scoring system to classify variants in different genes more precisely(8). This concise approach is complex and in need of validation. (3)3. Further initiatives such as ClinVar, an archival database for genetic variants, have allowed laboratories to share and compare the classification of a concrete variant and to resolve discordant interpretations (9). (4)4. Another partner initiative to ClinVar, the Clinical Genome Resource (ClinGen)(10), incorporates disease- and gene-specific knowledge to define the clinical relevance of genes and variants. Up to date, very limited data on variant curation is available. Knowledge of disease- and gene-specific characteristics is extremely useful for variant classification and advocate for additional disease- and gene-specific guidelines(11).

The *FBN1* gene (*FBN1*, OMIM #134797) encodes the multidomain protein fibrillin-1 (figure 1). This protein is composed of 47 Endothelial Growth Factor (EGF)-like domains of which 43 are calcium binding (cb-EGF), 7 TGFβ-binding (TB) domains, 2 hybrid (Hyb) domains and the N- and C-terminal domains. Each cb-EGF domain contains 6 cysteines, which are pairwise connected through disulphide bonds. Disruption of these bonds has been shown to render the protein more vulnerable to proteolysis and most likely affect its function(12–14). Furthermore, cb-EGF domains contain specific cb-sites ((D/N) X (D/N) (E/Q) X_m (D/N) X_n (Y/F) where m and n are variable number of residues and D: Aspartic acid, N: Asparagine E: Glutamic acid, Q: Glutamine, Y: Tyrosine and F:

Phenylalanine). These cb-sites confer structural stability to the protein, provide protection against degradation and control interaction with other components of the extracellular matrix(14,15).

Pathogenic variants in *FBN1* cause a range of connective tissue disorders, collectively known as type 1 fibrillinopathies(16). The most well-known and the focus of our study is Marfan syndrome (MFS, OMIM #154700, ORPHA #284963). MFS is an autosomal dominant inherited connective tissue disorder affecting multiple organ systems. Prominently affected are the cardiovascular, ocular, and skeletal systems, although manifestations in the skin, lungs and dura are also commonly associated. The diagnosis is made based on the revised Ghent criteria(17). In the index person and in the absence of family history, the combination of aortic root dilatation (ARD) or dissection and ectopia lentis (EL) or the combination of ARD or dissection and a positive systemic score will establish the diagnosis. Molecular analysis of the *FBN1* gene is not strictly necessary for the diagnosis but the identification of a pathogenic *FBN1* variant is helpful in those patients presenting with partial features of MFS. Moreover, it can also help to distinguish MFS from other related disorders presenting overlapping features such as Loeys-Dietz syndrome (LDS, [OMIM #609192 #610168, ORPHA #60030]), MASS phenotype (Mitral valve, Myopia, Aorta, Skin, Skeletal features [OMIM #604308, ORPHA #99715]) or mitral valve prolapse syndrome (MVPS, [OMIM #157700, ORPHA #741])(17).

This study addresses gene- and protein-specific characteristics in the interpretation of *FBN1* variants. Although some of the proposed refinements may be applicable to other genes, the majority are specific to the *FBN1* gene and to Marfan syndrome.

MATERIALS AND METHODS

The data, analytic methods, and study materials will be made available to other researchers for purposes of reproducing the results or replicating the procedure. The final classification of these variants will be published in the Clinvar database(18).

Sample selection

All consecutively identified variants from our in-house disease-specific MFS database classified as causal mutation or variant of uncertain significance between June 1990 and December 2016 were used for the study. A total of 713 *FBN1* variants were identified. With increasing knowledge, refinement of genetic techniques and collection of larger numbers of variants, variant classification inherently underwent modifications over this 26-year time period. In the past, classification was mainly based on the assessment of a combination of criteria, including whether or not the variant was located in an important functional domain, presence or absence of the variant in an in-house control population, co-segregation of the variant in the family, interspecies conservation of the affected nucleotide/amino acid and use of computational prediction programs (17)..

These 713 variants were found in a total of 934 probands referred for genetic testing in the context of a clinical suspicion for MFS and related disorders, EL syndrome, or familial or isolated thoracic aortic disease (TAD). Information on phenotypic features was available in 671 of the 934 probands. 201 underwent thorough clinical evaluation at our institution. In the remainder, phenotyping was performed by the referring physician and clinical data was extracted from clinical checklists or request forms.

Variant detection

FBN1 variants were identified using different mutation-detection methods, including single stranded conformation polymorphism (SSCP) and conformation sensitive gel electrophoresis (CSGE), denaturing high-performance liquid chromatography (dHPLC) analysis, Sanger sequencing, and next generation sequencing techniques. Some variants were only detected after using different techniques. All variants were confirmed with Sanger sequencing. Most of the analyses were performed on genomic DNA (gDNA) extracted from peripheral blood leukocytes. A small proportion

of the variants was identified in fibroblast-derived complementary DNA (cDNA). In case of gDNA sequencing, all coding exons and flanking intronic sequences (up to 20 base pairs) were analyzed.

Variant classification

In a working group consisting of molecular laboratory experts (WS, SS, PC, MR) and clinicians (LMM, BC, JDB) familiar with MFS and related disorders, we refined 15 of the 28 ACMG/AMP criteria to tailor them to the *FBN1* gene. The *FBN1* gene-specific refinements are summarized in table 1 and explained below. The genomic sequence used in the classification of the variants was the GRCh37 Human reference assembly (hg19).

Use of population databases

Population databases are used to determine the frequency of a certain variant in large populations. Common variants in these databases are more likely considered to be benign whereas absence of a variant in these population databases is usually regarded as a pathogenic criterium (PM2). In our study we considered all *FBN1* variants identified from the dbSNP(19), Exome Aggregation Consortium(20), Exome Variant Server(21), 1000Genome project(22) and the Genome of the Netherlands(23) databases. Since MFS is a rare disorder with an estimated prevalence between 1.5-17.2 per 100.000 individuals(24), the ACMG/AMP allele frequency necessary to be considered as a stand-alone criterium pro benignity (BA1) was reduced from 5 to 1%. Furthermore, an allele frequency above 0.0002 (1:5000) was considered as a strong criterium for benignity (BS1). Finally, case-control studies were only considered to use as a strong criterium (PS4) if more than 1000 controls were tested.

Use of phenotypic data and alternate locus observations

A patient's phenotype or family history highly specific for a monogenic disease is considered supportive evidence for pathogenicity (PP4). In case of MFS, clinical diagnosis can be established either if a patient presents ARD or dissection and EL or ARD or dissection and a positive systemic score. Different studies showed, however, that other syndromic forms of TAD(17,25) phenotypically overlap with MFS and can present not only ARD but also systemic features similar to MFS. The combination of ARD or dissection and EL seems to be more specific for MFS (17) and we therefore only applied the PP4 criterium when a patient presented these two features combined. In addition, when a (likely) pathogenic ((LP) variant was identified in another known TAD gene(26) it supported the notion that the *FBN1* variant was possibly benign (BP5).

Use of literature and (locus-specific) databases

Variant interpretation published in literature or (locus-specific) databases can be used in support of the classification of a variant of interest (PP5 or BP6). In our study we used the Human Gene Mutation Database (HGMD)(27), ClinVar(18), PubMed(28), Leiden Open Variation Database (LOVD)(29) and the Universal Mutation Database for *FBN1* (UMD-FBN1)(30) databases to search for a specific variant. Only publications providing sufficient clinical or functional evidence were considered for the study. Papers published by our own group were not taken into account.

Use of computational and predictive data

This category includes a variety of computational and predictive criteria, such as the outcome of *in silico* software tools that predict the impact of a certain variant (BP4, PP3), the predicted effect on the protein's length or expression (BP3, BP7, PM4) and previously reported pathogenic variants affecting the same amino acid or nucleotide (PM5, PS1). To evaluate computational data we used the following programs: MutationTaster(31), SIFT(32), PolyPhen2(33), Align GVGD(34) and Grantham(35) for the missense variants, and Human Splicing Finder(36), GeneSplicer(37), NNSplice(38) and MaxEntScan(39) for the splice site variants. A variant was considered to have a

predicted deleterious effect (PP3) or no effect (BP4) on the gene or gene product when at least 3 of the prediction models supported pathogenicity or benignity. If a novel missense variant affected the same nucleotide or amino acid that was previously reported to be (L)P, we considered this criterion as strong (PS1) or moderate (PM5) only when the reported variant was (L)P according to the ACMG/AMP guidelines and if well-established functional studies showed a deleterious effect of that particular variant. Since loss of expression or haploinsufficiency is known as a disease-causing mechanism in the *FBN1* gene, the PSV1 criterion was used for all frameshift and nonsense variants not affecting the final exon 65 and for all splice site variants in positions $\pm 1-2$. Frameshift and nonsense variants affecting exon 65 usually produce a protein which is shorter(2) and therefore we used for these cases the PM4 criterion.

Use of functional data

Robust functional studies showing an effect of the variant on mRNA or protein level are considered strong evidence for pathogenicity (PS3) or benignity (BS3). Cysteine substitutions in the cb-EGF domains of fibrillin-1 are examples of variants affecting well-established and important functional domains with multiple reports of detrimental functional consequences(13,14). This type of variant was considered as a strong criterion of pathogenicity. Variants affecting cysteine residues outside cb-EGF domains or residues within the highly conserved cb-sites sequences are less well established and were regarded as moderate evidence of pathogenicity (PM1)(14,15).

Use of segregation data

The ACMG/AMP guidelines contain several criteria for evaluating segregation. The PM6 can be used if a *de novo* variant is identified. When paternity and maternity is confirmed, the level of evidence can be increased to a strong criterion (PS2). The guidelines only contemplate one criterion for co-segregation of the variant with the phenotype irrespective to the number of affected family members carrying the variant (PP1). We adapted the guidelines according to the

number of affected family members in whom the variant was present, with 1-2 family members affected being a supportive criterium (PP1), 3-4 family members being a moderate criterium (PM7) and >4 family members being a strong criterium (PS5).

Statistical analysis

Statistical analysis was performed using IBM SPSS Statistics 24 package (SPSS Inc., Chicago, IL, USA). Unless stated otherwise, continuous variables are expressed as mean and standard deviation and categoric variables as absolute value and percentage. χ^2 and the McNemar tests were used to analyze categorical variables and the K-alpha test to evaluate the agreement between the new classification and the old practices. The Kolmogorov-Smirnov test was used to determine normality. Normal distributed variables were analyzed using the unpaired sample t-test and the ANOVA and non-normal distributed variables were analyzed using the Mann-Whitney-U and the Kruskal Wallis tests. A p value of < 0.05 was used to define statistical significance (two-sided).

The study was approved by the local Independent Ethics Committee (IEC) and the Institutional Review Board (IRB) of our hospital.

RESULTS

A total of 713 *FBN1* variants have been identified in our lab in 934 unrelated probands (358 [53.4 %] male, mean age 25.7 ± 16.4 yrs, 42.8% under the age of 20 yrs). The majority of these variants were missense variants (n=414, 58.1%), followed by frameshift (n=118, 16.5%), nonsense (n=81, 11.4%), splice site (n=80, 11.2%), inframe (n=12, 1.7%) and synonymous (n=8, 1.1%) variants. Of the missense variants, 159 (38.4%) affected a cysteine residue in a cb-EGF domain and 104 (25.1%) affected either a cysteine residue outside a cb-EGF domain or an amino acid in a cb-site. Of the splice site variants, 43 (53.75%) were in a ± 1 or 2 position (figure 2).

Overall agreement between the new adapted guidelines and the old practices was 86.4% (K-alpha 0.6). The majority of the variants were reclassified as (L)P (n=526, 73.8%), this was significantly lower in comparison to old practices in which the amount of variants classified as causal mutation was 610 (85.5%, $p < 0.001$). Consequently, the overall amount of VUS increased from 14.5% to 24.2%. Only 14 variants (2%) were downscaled to LB (figure 3) and 1 variant was upscaled from VUS to LP. The great majority of the variants in which the new and old interpretation disagreed were downscaled from causal mutation to LB (3) or VUS (82) (table 2). A detailed description of the variants in which there was disagreement can be found in supplemental table 1. The majority of the downscaled variants from causal mutation to LB or VUS were missense variants outside a cb-EGF domain and not affecting a cysteine residue (n=49, 59.7%), followed by missense variants affecting either a cysteine residue outside a cb-EGF domain or an amino acid in a cb-site (n=13, 15.1%), splice site variants outside ± 1 or 2 position (n=7, 8.1%), frameshift variants in exon 65 (n=7, 8.1%), inframe variants (n=4, 4.7%), synonymous variants (n=3, 3.5%) and nonsense variants in exon 65 (n=2, 2.3%) (table 3). The clinical data needed to correlate phenotype-genotype and segregation analysis was (partly) unavailable in almost half of the cases in which a variant was downscaled (n=35, 40.7% and n=41, 47.7% respectively).

All frameshift and nonsense variants not affecting exon 65, all splice site variants in positions $\pm 1-2$ and all missense variants affecting a cysteine residue in a cb-EGF domain were classified as (L)P (figure 4). Only one of the 8 frameshift and one of the 3 nonsense variants affecting exon 65 could be classified as LP, the rest were classified as VUS. Of the 104 missense variants affecting either a cysteine residue outside a cb-EGF domain or an amino acid in a cb-site, 90 (86.5%) were classified as (L)P. This percentage was much lower for other types of variants: only 32 (21.2%) of the remaining missense variants and 7 (58.3%), 2 (25%) and 6 (16.2%) of the inframe, synonymous and splice site variants outside the positions $\pm 1-2$, respectively, were (L)P (figure 4).

Clinical data were available in 671 of the 931 cases. Two hundred sixty-two probands (39%) fulfilled the revised Ghent criteria based on clinical data alone, of whom 160 (61.06%) had ARD or dissection and EL. The remaining had ARD or dissection and a positive systemic score. Of the individuals fulfilling the revised Ghent criteria, 228 (87.1%) had a (L)P *FBN1* variant. This percentage was significantly higher in comparison to the group not fulfilling the Ghent criteria (number of (L)P variants 288, 70.4%, $p < 0.001$). Of the 34 cases fulfilling the revised Ghent criteria but in whom a VUS or LB variant was identified, 5 individuals carried a frameshift variant in exon 65, 1 individual a nonsense variant in exon 65 and 1 individual a variant affecting a cysteine residue outside a cb-EGF domain. The 27 other patients had either a missense variant outside a well-established critical domain or a splice site variant outside positions $\pm 1-2$. In three of these 27 patients an additional variant of unknown significance was found: two in *FBN1* and one in *COL3A1*.

Patients not fulfilling the revised Ghent criteria were significantly younger than those who did (24.42 ± 17.09 versus 27.65 ± 15.23 , $p = 0.011$). This difference was even more apparent when comparing only the groups in which a (L)P variant was found (21.96 ± 16.09 versus 27.35 ± 15.30 , $p < 0.001$) illustrating the age-dependent expression of MFS.

As mentioned in the 'materials and methods' section, a distinction in the level of pathogenicity was made based on the amount of family members in which a certain variant in *FBN1* segregated with the phenotype. The moderate (PM7) and the strong (PS5) criteria were used in 20 and in 6 of the classified variants respectively. In only 3 of these cases a VUS was upscaled to a LP variant. In eighty-six cases, the *de novo* character of the variant was confirmed. In 76 (88.4%) of these 86 cases the variant was classified as (L)P. In 21 (27.6%) of these 76 cases de variant was upscaled from VUS to LP.

DISCUSSION

The publication of the ACMG/AMP guidelines was an important step forward in the classification of genetic variants. However, gene specific criteria were not incorporated in these guidelines allowing for subjective interpretation and considerable inter-observer variability in classification outcomes(3).

In our study we tailored some of the ACMG/AMP criteria to specifically suit *FBN1* variant interpretation. When applying the tailored criteria to our cohort, we downscaled 13.9% of the causal mutations to VUS or LB (table 2) and upscaled only one from VUS to LP. The total percentage of (L)P variants found with the new classification was significantly lower in comparison to the previously used classification methods (73.8% versus 85.5%, $p < 0.001$). This difference can be explained by several factors. First and most important is the increased stringency of the pathogenicity criteria in the new ACMG/AMP guidelines. For example, a missense variant that did not affect an important functional site but segregated in the family and was absent in 400 ethnically matched controls would previously have been considered pathogenic. According to the new guidelines, however, a higher level of evidence is necessary to classify this variant as pathogenic (absence from population databases, clear phenotype-genotype correlation, computational data which predicts pathogenicity and/or functional studies demonstrating that the variant is pathogenic).

Second, in the new ACMG/AMP guidelines the variant type strongly influences the final classification. Often, null variants are likely disrupting gene product and function, therefore counting as a very strong criterium (PVS1) in the variant classification. Similarly, variants affecting a well-established *in vitro* or *in vivo* functional site (PS3) are likely to be classified as (L)P. In contrast, other variant types require additional careful phenotyping, segregation analysis, computational tools and, in some cases, functional studies in order to enable classification as either benign or pathogenic. This phenomenon is clearly reflected in the results of our study. All frameshift and

nonsense variants not affecting exon 65, all splice-site variants at positions $\pm 1-2$ and all variants affecting a cysteine residue in a cb-EGF domain were classified as (L)P whereas a variable percentage of (L)P variants was found in the other variant types (figure 4). Consequently, it is not surprising that the majority of the variants downscaled from causal to VUS or LB were in fact missense variants not affecting a well-established functional site (59.7%). The combination of this high percentage of missense variants, insufficient phenotypical or segregation data (40.7% and 47.7% of the downscaled variants respectively) and stricter guidelines, can partially explain the difference found in our study between the old and the new classification systems.

Besides the general increase in stringency of the new ACMG/AMP guidelines, some of the further refinements applied in our study might also have led to the difference in outcome. Some of the refined criteria, which are open for further debate, are discussed here. First, the allele frequency cut-off for the BS1 criterium was established based on the prevalence of MFS alone, however, a more robust statistical analysis to determine which frequency of a variant in a reference sample is acceptable might be necessary. A good example is the recent publication of Whiffin et al.(40) in which not only the prevalence of the disease is taken into consideration, but also the estimated contribution of the gene and allele to the disease. Second, the PP4 criterium was used if a patient had ARD or dissection and EL, a combination of features considered more specific for MFS. However, some laboratories might also consider the use of the PP4 criterium if a patient presents a combination of ARD or dissection and a positive systemic score. In our study the classification of 5 patients would have changed from VUS to LP if the latter consideration had been used. Third, the analysis of data coming from different exome sequencing projects has led to the realization that many variants earlier thought to be pathogenic, are actually present in healthy individuals and are therefore more likely to be considered rare benign variants. Taking this into consideration, in our study the PP5 criterium was used only when sufficient data was provided in the original publication to account the variant as truly pathogenic. Furthermore, since the PM5 and PS1 criteria (variant

previously published affecting the same amino acid or the same nucleotide) weigh stronger on the final classification of a variant, these were only used if additional functional evidence of pathogenicity was available. Fourth, while most would agree that variants in *FBN1* affecting a cysteine residue in a cb-EGF domain should be considered as a strong criterium of pathogenicity (PS3), less consensus may exist as to deem the cysteine substitutions in other domains or the amino acid substitutions affecting the cb-sites as moderate or strong criterium pro pathogenicity. We decided to make a distinction between these two categories based on the number of functional reports. Multiple studies have been published showing that cysteine substitutions in different cb-EGF domains have a deleterious effect on the protein(13,14,41) and therefore, we decided to generalize this to all cysteine substitutions in these particular domains. In contrast, only 9 studies reported a deleterious effect of missense variants in conserved cb-sites (most affecting a different amino acid of the consensus sequence) and only 2 studies reported on a deleterious effect of cysteine residues in a hybrid or TB domain in fibrillin-1(14,15). This was judged insufficient to use as a general rule to support a strong criterium. Fifth, we applied an arbitrary cut-off to the co-segregation rule (PP1), making it a supportive criterium in case 1 or 2 affected family members carried the variant, a moderate criterium when 3-4 affected family members carried the variant, and a strong criterium when the number of affected family members carrying the variant was above 4. This rule may, however, be in need of more objective and statistic support. Contrary to the previous refinements, this refinement actually helped classifying 3 variants as LP whereas otherwise they would have been VUS.

During the study period we came to realize that some highly relevant criteria were not included in the current guidelines. First, the frequency with which a particular variant has been reported by independent laboratories could also be considered as a criterium pro pathogenicity. In this case, statistical analysis to determine cut-off values (number of reports) for the strength of the argument would be required. Second, regarding co-segregation analysis, currently only affected patients

carrying the variant or *de novo* cases count for the classification. Healthy non-carriers and obligate carriers within a family in whom genetic testing is not possible, could, however, also be taken into account when studying co-segregation. Since expression of MFS , especially in young individuals, can be variable, healthy non- carriers should have benefited from careful clinical, echocardiographic and ophthalmologic examination before being considered for the classification.

In conclusion, variant classification is a dynamic process that can change over time. In this paper we have refined some of the criteria of the ACMG/AMP guidelines to tailor them to the specific characteristics of the *FBN1* gene and its protein in order to allow a more precise and consistent interpretation of *FBN1* variants. Some of the criteria proposed here might be in need of further debate. Sharing data on functional studies, careful annotation of the variants and appropriate and standard reporting of clinical features could aid variant classification tremendously and subsequently guide clinical decision-making adequately.

Financial and competing interest disclosure

Laura Muiño-Mosquera is supported by a doctoral fellowship from the Special Research Fund (BOF) of the Ghent University.

Felke Steijns and Ilse Meerschaut are supported by a Methusalem grant of the Flemish government and Ghent University to A. De Paepe.

Marjolijn Renard is a postdoctoral researcher of the Research Foundation Flanders.

Bert Callewaert and Julie De Backer are senior clinical investigators of the Research Foundation Flanders (FWO).

REFERENCES

1. Pepin MG, Murray ML, Bailey S, Leistriz-Kessler D, Schwarze U, Byers PH. The challenge of comprehensive and consistent sequence variant interpretation between clinical laboratories. *Genet Med*. 2016; 18 (1): 20-4.
2. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J. 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.
3. Amendola LM, Jarvik GP, Leo MC, McLaughlin HM, Akkari Y, Amaral MD, et al. Performance of ACMG-AMP Variant-Interpretation Guidelines among Nine Laboratories in the Clinical Sequencing Exploratory Research Consortium. *Am J Hum Genet*. 2016;98(6):1067–76.
4. Hoskinson DC, Dubuc AM, Mason-Suares H. The current state of clinical interpretation of sequence variants. *Curr Opin Genet Dev*. 2017;42:33–9.
5. Patel RY, Shah N, Jackson AR, Ghosh R, Pawliczek P, Paithankar S, et al. ClinGen Pathogenicity Calculator: a configurable system for assessing pathogenicity of genetic variants. *Genome Med*. 2017;9(1):3.
6. Li Q, Wang K. InterVar: Clinical Interpretation of Genetic Variants by the 2015 ACMG-AMP Guidelines. *Am J Hum Genet*. 2017;100(2):267–80.
7. Kleinberger J, Maloney KA, Pollin TI, Jeng LJ. An openly available online tool for implementing the ACMG/AMP standards and guidelines for the interpretation of sequence variants. *Genet Med*. 2016;18 (11): 1165.
8. Nykamp K, Anderson M, Powers M, Garcia J, Herrera B, Ho Y-Y, et al. Sherlock: a comprehensive refinement of the ACMG/AMP variant classification criteria. *Genet Med*. 2017; 19 (10): 1105-1117
9. Harrison SM, Dolinsky JS, Knight Johnson AE, Pesaran T, Azzariti DR, Bale S, et al. Clinical Laboratories Collaborate to Resolve Differences in Variant Interpretations Submitted to ClinVar. *J Am Coll Med Genet*. 2017;19(10):1096–104.
10. Rehm HL, Berg JS, Brooks LD, Bustamante CD, Evans JP, Landrum MJ. ClinGen--the clinical genome resource. *N Engl J Med*. 2015;372 (23): 2235-42.
11. Vincent L. Optimization and utilization of ACMG variant classification criteria for the RASopathies: A ClinGen initiative. In: *American Society of Human Genetics*. 2015. Abstract
12. Schrijver I, Liu W, Odom R, Brenn T, Oefner P, Furthmayr H, et al. Premature Termination Mutations in FBN1: Distinct Effects on Differential Allelic Expression and on Protein and Clinical Phenotypes. *Am J Hum Genet*. 71(2):223–37.
13. Vollbrandt T, Tiedemann K, El-Hallous E, Lin G, Brinckmann J, John H, et al. Consequences of Cysteine Mutations in Calcium-binding Epidermal Growth Factor Modules of Fibrillin-1. *J Biol Chem*. 2004; 279(31):32924–31.

14. Zeyer KA, Reinhardt DP. Engineered mutations in fibrillin-1 leading to Marfan syndrome act at the protein, cellular and organismal levels. *Mutat Res Mutat Res*. 2015;765:7–18.
15. Reinhardt DP, Ono RN, Notbohm H, Müller PK, Böhlinger HP, Sakai LY. Mutations in Calcium-binding Epidermal Growth Factor Modules Render Fibrillin-1 Susceptible to Proteolysis: A potential disease-causing mechanism in Marfan syndrome. *J Biol Chem*. 2000; 275(16):12339–45.
16. Sakai LY, Keene DR, Renard M, De Backer J. FBN1: The disease-causing gene for Marfan syndrome and other genetic disorders. *Gene*. 2016; 591(1):279–91.
17. Loeys BL, Dietz HC, Braverman AC, Callewaert BL, De Backer J, Devereux RB, et al. The revised Ghent nosology for the Marfan syndrome. *J Med Genet*. 2010; 47(7):476–85.
18. Landrum MJ, Lee JM, Benson M, Brown G, Chao C, Chitipiralla S. ClinVar: public archive of interpretations of clinically relevant variants. *Nucleic Acids Res*. 2016;44 (D1):D862-8
19. Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, Smigielski EM. dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res*. 2001;29.
20. Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T. Analysis of protein-coding genetic variation in 60,706 humans. *Nature*. 2016;536 (7616): 285-91.
21. Exome variant server, NHLBI GO Exome Sequencing Project (ESP). Seattle WA. Available from: <http://evs.gs.washington.edu/EVS/>
22. The 1000 Genomes Project Consortium. A global reference for human genetic variation. *Nature*. 2015 Oct 1;526(7571):68–74.
23. The Genome of the Netherlands Consortium. Whole-genome sequence variation, population structure and demographic history of the Dutch population. *Nat Genet*. 2014; 46(8):818–25.
24. von Kodolitsch Y, De Backer J, Schüler H, Bannas P, Behzadi C, Bernhardt AM, et al. Perspectives on the revised Ghent criteria for the diagnosis of Marfan syndrome. *Appl Clin Genet*. 2015;8:137–55.
25. Campens L, Callewaert B, Muiño Mosquera L, Renard M, Symoens S, De Paepe A, et al. Gene panel sequencing in heritable thoracic aortic disorders and related entities – results of comprehensive testing in a cohort of 264 patients. *Orphanet J Rare Dis*. 2015; 10 (1).
26. Milewicz DM, Prakash SK, Ramirez F. Therapeutics Targeting Drivers of Thoracic Aortic Aneurysms and Acute Aortic Dissections: Insights from Predisposing Genes and Mouse Models. *Annu Rev Med*. 2017;68(1):51–67.
27. Stenson PD, Mort M, Ball EV, Evans K, Hayden M, Heywood S, et al. The Human Gene Mutation Database: towards a comprehensive repository of inherited mutation data for medical research, genetic diagnosis and next-generation sequencing studies. *Hum Genet*. 2017;136(6):665–77.
28. NCBI Resource Coordinators. Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res*. 2016;44:D7–19.

29. Fokkema IFAC, Taschner PEM, Schaafsma GCP, Celli J, Laros JFJ, den Dunnen JT. LOVD v.2.0: the next generation in gene variant databases. *Hum Mutat.* 2011;32(5):557–63.
30. Collod-Bérout G, Le Bourdelles S, Ades L, Ala-Kokko L, Booms P, Boxer M, et al. Update of the UMD-FBN1 mutation database and creation of an FBN1 polymorphism database. *Hum Mutat.* 2003;22(3):199–208.
31. Schwarz JM, Cooper DN, Schuelke M, Seelow D. MutationTaster2: mutation prediction for the deep-sequencing age. *Nat Meth.* 2014;11(4):361–2.
32. Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat Protoc.* 2009;4(8):1073–81.
33. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, et al. A method and server for predicting damaging missense mutations. *Nat Methods.* 2010;7(4):248–9.
34. Tavtigian SV, Deffenbaugh AM, Yin L, Judkins T, Scholl T, Samollow PB, et al. Comprehensive statistical study of 452 BRCA1 missense substitutions with classification of eight recurrent substitutions as neutral. *J Med Genet.* 2006;43(4):295.
35. Grantham R. Amino Acid Difference Formula to Help Explain Protein Evolution. *Science.* 1974; 185(4154):862.
36. Desmet F-O, Hamroun D, Lalande M, Collod-Bérout G, Claustres M, Bérout C. Human Splicing Finder: an online bioinformatics tool to predict splicing signals. *Nucleic Acids Res.* 2009; 37(9):e67–e67.
37. Pertea M, Lin X, Salzberg SL. GeneSplicer: a new computational method for splice site prediction. *Nucleic Acids Res.* 2001; 29(5):1185–90.
38. Reese MG, Eeckman FH, Kulp D, Haussler D. Improved Splice Site Detection in Genie. *J Comput Biol.* 1997;4(3):311–23.
39. 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.
40. Whiffin N, Minikel E, Walsh R, O’Donnell-Luria AH, Karczewski K, Ing AY, et al. Using high-resolution variant frequencies to empower clinical genome interpretation. *Genet Med.* 2017;19(10):1151–8.
41. Schrijver I, Liu W, Brenn T, Furthmayr H, Francke U. Cysteine substitutions in epidermal growth factor–like domains of fibrillin-1: distinct effects on biochemical and clinical phenotypes. *Am J Hum Genet.* 1999;65(4):1007–1020.

Table 1:Criteria used for the reclassification of the variants in *FBN1*

Categories	Resource	Criteria of the ACMG/AMP guidelines	Specifications of the <i>FBN1</i> classification
Population data	dbSNP ExAC Exome Variant Server 1000 Genome project GoNL	BA1 allele frequency is >5% BS1 allele frequency is greater than expected for the disorder BS2 Observed in a healthy adult individual with full penetrance expected at an early age PM2 variant is absent from controls PS4 the prevalence of the variant in affected individuals is greater than in controls	BA1 allele frequency is >1% BS1 allele frequency is > 0.0002 (1:5000) PS4 this criterium was used only if at least 1000 controls were included
Genotype and phenotype data		BP5 patient has an alternate molecular basis for the disease* PP4 patient’s phenotype or family history is highly specific for a disease with a single genetic origin	 PP4 patient has TAD and EL

<p>Literature and (locus specific) database</p>	<p>HGMD ClinVar PubMed LOVD UMD-FBN1</p>	<p>BP6 variant reported as benign PP5 variant reported as pathogenic</p>	<p>BP6 and PP5 were used only after critical review of the literature</p>
<p>Computational and predictive data</p>	<p><i>Missense variants</i> Mutation taster SIFT PolyPhen-2 Align GVGD Grantham <i>Splice site variants</i> Human Splicer Finder GeneSplicer NNSplice</p>	<p>BP1 missense variant in a gene for which primarily truncating variants are known to cause disease BP3 inframe deletions/insertions in a repetitive region without a known function. BP4 multiple lines of computation evidence suggest no impact on gene or gene product. BP7 synonymous variant for which splicing prediction algorithms predict no impact PP3 multiple lines of computation evidence suggest impact on gene or gene product. PM4 protein length changes as a result of inframe</p>	<p>BP4 at least 3 lines of computation evidence suggest no impact on <i>FBN1</i> or production of fibrillin-1 PP3 at least 3 lines of computation evidence suggest impact on <i>FBN1</i> or production of fibrillin-1 PM4 protein length changes as a result of inframe change or frameshift or nonsense variant affecting exon 65</p>

	MaxEntScan	<p>deletion or insertion</p> <p>PM5 novel missense change at an amino acid residue where a different pathogenic missense change has been seen before</p> <p>PS1 same amino acid change previously established as pathogenic regardless of nucleotide change</p> <p>PVS1 frameshift, nonsense, canonical $\pm 1-2$ splice site or exon deletion.</p>	<p>PM5 and PS1 were used only after critical review of the literature and if functional data was available.</p> <p>PVS1 frameshift, nonsense, canonical $\pm 1-2$ splice site or exon deletion (except frameshift or nonsense affecting exon 65)</p>
Functional data	<p>HGMD</p> <p>ClinVar</p> <p>PubMed</p> <p>LOVD</p> <p>UMD-FBN1</p>	<p>BS3 well established in vitro or in vivo studies show no damaging effect</p> <p>PP2 Missense variant in a gene which has low rate of benign missense variation and which missense variants a well-known disease mechanism</p> <p>PM1 located in a mutational hot-spot and/or critical established functional domain</p>	<p>PM1 located in a mutational hot-spot and/or critical established functional domain</p> <p><i>Cys substitutions outside cb-EGF domain</i></p>

		<p>PS3 well established in vitro or in vivo studies show damaging effect</p>	<p><i>Introduction of a new Cys within cb-EGF domain</i></p> <p><i>Cb-site substitutions in cb-EGF domain</i></p> <p>PS3 well established in vitro or in vivo studies show damaging effect</p> <p><i>Cys substitution within cb-EGF domain</i></p>
Segregation data	Segregation analysis	<p>BS4 lack of segregation in affected members of the family</p> <p>PP1 co-segregation with the disease in multiple affected family members in a gene known to cause the disease</p> <p>PM6 assumed <i>de novo</i>, without confirmation of paternity/maternity</p> <p>PS2 <i>de novo</i> (paternity/maternity confirmed) in a</p>	<p>PP1 co-segregation with the disease in 1-2 family members</p> <p>PM7 co-segregation with the disease in 3-4 family members</p> <p>PS5 co-segregation with the disease in >4 family members</p>

		patient with the disease and no family history	
Allelic data		<p>PM3 For recessive disorders detected in <i>trans</i> with a pathogenic variant</p> <p>BP2 Observed in <i>trans</i> with a pathogenic variant for a fully penetrant dominant gene or in <i>cis</i> with a pathogenic variant in any inheritance pattern</p>	

*Other TAD genes considered: TGFBR1/2, SMAD3, ACTA2, TGFB2 and COL3A1

Coding system: Pathogenic criteria- PVS: very strong, PS: strong, PM: moderate, PP: supportive- Benign criteria- BA: stand alone, BS: very strong, BM: moderate, PP: supportive

Abbreviations: ACMG: American College of Medical Genetics, AMP: Association for Molecular Pathology, ExAC: Exome Aggregation Consortium, GoNL: Genome of the Netherlands, HGMD: Human Gene Mutation Database, LOVD: Leiden Open Variation Database, TAD: thoracic aorta disease, EL: ectopia lentis, Cys: cysteine, cb-EGF: calcium binding- epidermal growth factor, cb-site: calcium binding site.

Table 2: Comparison of the old and new classification

		New Classification based on the ACMG/AMP guidelines				
		LB (%)	VUS (%)	LP (%)	P (%)	Total
Old classification	VUS (%)	11 (10.7)	91 (88.3)	1 (1)	0	103
	pathogenic variant (%)	3 (0.5)	82 (13.4)	353 (57.9)	172 (28.2)	610
	Total	14	173	354	172	713

Abbreviations: LB: likely benign, VUS: variant of unknown significance, LP: likely pathogenic, P: pathogenic

Table 3: Characteristics of the downscaled variants

	Characteristics of the downscaled variants	
Type of variant	Frameshift	7 (8.1)
N (%)	Nonsense	2 (2.3)
	Cys outside cb-EGF or cb-site	13 (15.1)
	Other missense	49 (59.7)
	Splice-site non \pm 1-2	7 (8.1)
	Inframe	4 (4.7)
	Synonymous	3 (3.5)
	Ghent criteria	Yes
N (%)	No	29 (35.4)
	Insufficient data	35 (40.7)
Family history	Negative	24 (27.9)
N (%)	Proven de novo (no maternity/paternity tested)	4 (4.7)
	Positive	16 (19.5)
	Insufficient data	41 (47.7)

*Seven of these patients presented ectopia lentis and aortic root dilation

Figure legends

Figure 1: Schematic representation of the FBN1 protein and the cb-EGF domains

Figure 2: Type of variants found in *FBN1*

Abbreviations: Cys: cysteine, cb-EGF: calcium binding Epithelial Growth Factor domain, cb-site: calcium binding site

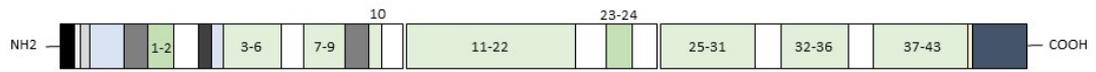
Figure 3: Degree of agreement between the new adapted guidelines to *FBN1* and old practices

Abbreviations: LB: likely benign, VUS: variant of uncertain significance, (L)P: (likely) pathogenic

Figure 4: Amount of identified variants per level of pathogenicity

Fig. 2: A. absolute count and percentage of variants in each level of pathogenicity. B. Representation of the level of pathogenicity per type of variant.

Figure 1: Schematic representation of the FBN1 protein and the cb-EGF domains



cb-EGF-like	p-position	exon
1-2	246-329	7-8
3-6	490-653	12-15
7-9	723-846	18-20
10	910-951	23
11-22	1028-1527	25-36
23-24	1606-1688	39-40
25-31	1766-2054	43-50
32-36	2127-2332	52-56
37-43	2402-2687	58-63

- Signal Peptide
- EGF-like
- Cb EGF-like
- NH2 unique region
- LTBP-like
- Proline-rich
- TGFβ binding
- Hybrid module
- COOH unique region
- Fibulin-like

Figure 2: Type of variants found in *FBN1*

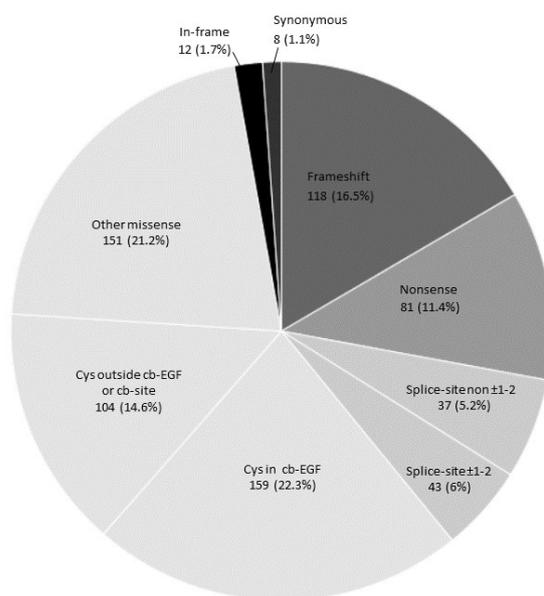


Figure 3: Degree of agreement between the new adapted guidelines to *FBN1* and the old practices

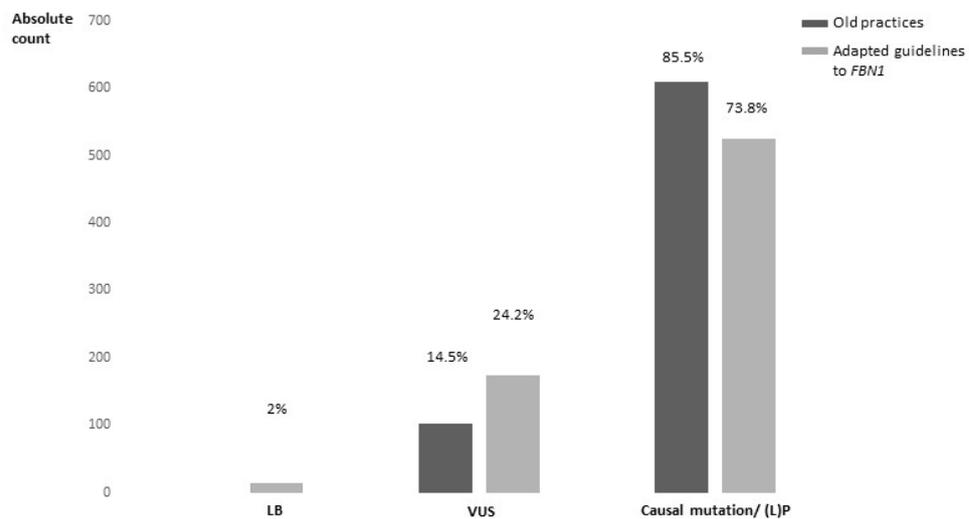


Figure 4: Amount of identified variants per level of pathogenicity

