A Reference Meta-Model to Understand DNA Variant Interpretation Guidelines

Mireia Costa[1,⋆][0000−0002−8614−0914], Alberto García S.[1,⋆][0000−0001−5910−4363], Ana Leon[1][0000−0002−0877−7063], Anna Bernasconi[2][0000−0001−8016−5750], and Oscar Pastor[1][0000−0002−1320−8471]

[1] PROS Group, Valencian Research Institute (VRAIN)
Universitat Politècnica de València
micossan@vrain.upv.es, algarsi3@pros.upv.es,
aleon@pros.upv.es,opastor@pros.upv.es
[2] Politecnico di Milano, Milan, Italy
anna.bernasconi@polimi.it

Abstract. Determining the role of a DNA variant in patients' health status – a process known as *variant interpretation* – is highly critical for precision medicine applications. Variant interpretation involves a complex process where, regrettably, there is still debate on how to combine and weigh diverse available evidence to achieve proper and consistent answers. Indeed, at the time of writing, 22 different variant interpretation guidelines are available to the scientific community, each of them attempting to establish a framework for standardizing the interpretation process. However, these guidelines are qualitative and vague by nature, which hinders their streamlined application and potential automation. Consequently, more precise definitions are needed. Conceptual modeling provides the means to bring clarification within this domain. This paper presents our efforts to define and use a UML meta-model that describes the main concepts involved in the definition of variant interpretation guidelines and the constructs they evaluate. The precise conceptual definition of the guidelines allowed us to identify four common misinterpretation patterns that hamper the correct interpretation process and that can consequently affect classification results. In several proposed examples, the use of the meta-model provides support in identifying the inconsistencies in the observed process; this result paves the way for further proposing reconciliation strategies for the existing guidelines.

Keywords: Conceptual Modeling · Genomics · Variant Interpretation Guidelines · Standards

1 Introduction

Precision medicine has emerged as a disruptive medical approach aiming to transform historically reactive medicine into a proactive one. To do so, this

⋆ M.C. and A.G.S. should be regarded as Joint First Authors.

new perspective prioritizes individualized clinical actions based on each patient's unique characteristics [53]. The most distinguishing characteristic of an individual is its DNA sequence, which slightly differs among individuals.

Individual DNA sequences are compared to a DNA reference sequence that reflects an "ideal" individual, leading to the identification of differences. These differences among individuals are known as DNA variants[3], and they determine our physical characteristics, predisposition to disorders, or a different response to treatments.

Identifying variants in an individual's DNA sequence has become easier and faster thanks to Next-Generation Sequencing (NGS) [41]. This technique uses massive parallelization to obtain the entire DNA sequence of an individual; the connected technological advancement has significantly improved our ability to identify and analyze DNA variants [41]. However, the scientific community must overcome numerous challenges (costs, ethics, security of the shared data, and data integration and interpretation, among others) [34] before achieving the paradigm shift that precision medicine proposes. In the data integration and interpretation context, one of the most difficult challenges is determining a DNA variant's role in our health status (i.e., whether it will cause a particular disorder or affect treatment response), a process known as *variant interpretation*.

Variant interpretation is a complex process that involves weighing various factors, such as the variant's frequency among the population, whether it has previously been linked to a disorder, etc. Geneticists and clinical experts are still debating on how to correctly weigh this evidence in order to achieve proper variant interpretation. To address this issue, several authors have developed *variant interpretation guidelines*. A variant interpretation guideline is a set of instructions designed to guide the interpretation process by assessing whether or not a variant meets specific criteria. These guidelines have quickly been embraced by geneticists [37] and they have been adapted to the peculiarities of several disorder-causing genes [30].

However, several issues have arisen due to the vague definition of these guidelines and their application, which depends on the subjective interpretations of domain experts [50]. In this context, clinical experts argue that more concrete definitions are needed to standardize the variant interpretation process and reduce inconsistencies [2]. A suitable approach to clarify this complex domain is Conceptual Modeling. Conceptual Modeling techniques have proven to be effective to achieve high levels of concreteness and standardization in genomics [42,38,45,7,15].

In this work, we report on our use of Conceptual Modeling to achieve a systematized definition of the main concepts involved in the definition of variant interpretation guidelines and the constructs they evaluate. For this purpose, we created a meta-model using the Unified Modeling Language [10]. Twenty-two well-known variant interpretation guidelines were carefully analyzed. Based on these analyses, we were able to characterize the differences and similarities between these guidelines. This allowed us to identify the common conceptual

[3] https://www.cancer.gov/publications/dictionaries/genetics-dictionary/def/variant

structure that underpins all of these guidelines and to consolidate our findings via the meta-model.

The contribution of our work is to show how using a conceptual meta-model to represent the concepts and constructs behind variant interpretation guidelines can provide the following benefits: (a) Definition of the underlying structure that different interpretation guidelines share, resulting in the development of a common framework for representing various types of guidelines; (b) Identification of patterns of misinterpretation of variants due to inconsistencies or conflicts within or between existing guidelines; (c) Disentanglement of the intricate details of existing clinical guidelines by resolving aspects whose definitions are left implicit or ambiguous, requiring clarification.

Prospectively, our contribution can support a shared effort to define clinical guidelines more consistently and objectively, reducing variant interpretation inconsistencies. In parallel, it offers the possibility to improve variant interpretation automation because tools will be based on a precise and concrete definition to guide their implementation rather than relying on personal interpretations.

The remainder of the paper is organized as follows. Section 2 provides the background that has motivated our work. Section 3 overviews related work. Section 4 describes the proposed conceptual meta-model, instantiating it on a simple example of use (contributing to benefit (a)). Section 5 proposes to use the above-mentioned conceptual meta-model to define a set of misinterpretation patterns (contributing to benefit (b)). Section 6 discusses lessons learned (regarding benefit (c)) and, finally, Section 7 concludes the paper with a future outlook.

2 Background

At the time of writing, 22 DNA variant interpretation guidelines have been proposed. Some examples are the ACMG/AMP 2015 guidelines [43], the Ambry Genetics 2015 guidelines [17], or the ACGS for Rare Disease 2020 guidelines [23]. The current guidelines support Mendelian disorders (disorders caused by variants in a single gene), Rare disorders (disorders affecting a small percentage of the population), X-linked disorders (disorders caused by variants in the X chromosome), and Recessive or Autosomal dominant disorders (disorders with a specific inheritance pattern). Some guidelines are only applicable to somatic variants (variants that occur after conception in specific body tissues), mitochondrial variants (variants affecting the mitochondrial DNA), or copy number variants (variants that affect the number of copies of a specific gene). Finally, some guidelines present generic applicability, i.e., they are theoretically applicable to any kind of variant or disorder.

Even though these guidelines attempted to improve and standardize the variant interpretation process, they are far from being a shared and widely-adopted solution. Indeed, distinct works [13,21] have highlighted a number of issues that arise when using variant interpretation guidelines. The most frequently expressed concern is that the guidelines are qualitative in nature without providing the

needed specificity [28,24]; as a consequence, their practical application is left open to expert interpretation [22]. In this context, inconsistent interpretations among experts become common, leading to serious consequences in healthcare applications.

Consider for instance a case where an initial assessment in prenatal care reveals that an unborn child is at high risk of developing Muscular Dystrophy disorder. The assessment was later revised by a different team of experts, finally determining that it was incorrect [40]. Because often these families have to make decisions on pregnancy management within a limited timeframe, the improperly classified variant could have had irreversible consequences. Furthermore, the more complex the disorder (e.g., cancer), the more inconsistencies in variant classification usually emerge [14].

In an effort to provide more exact definitions and streamline the process by lowering the complexity and time needed to complete the interpretation, several tools have been created to automate the variant interpretation process [26,29,35,46,52,49]. Among these, VarSome [26], InterVar [29], and CharGer [46] aim to operate within a broad scope, i.e., with variations associated with any kind of disorder. Instead, CardioVAI [35] and CardioClassifier [52] focuse on inherited cardiac conditions. All of them assign a label representing the disorder-causing potential of the variants based on a set of applied criteria from the ACMG/AMP 2015 guidelines. Following a different approach, Tavtigian et al. [49] modeled the ACMG/AMP 2015 guidelines as a Bayesian framework, which allowed the authors to provide a probabilistic score of pathogenicity associated with each variation.

These tools are meant to provide automated support for the variant interpretation process; this is supposed to be more effective than human application and reduce reproducibility issues. However, the qualitative nature and insufficient specificity of variant interpretation guidelines cause different tools to make assumptions and interpret the data in discordant ways. Furthermore, some guideline criteria are frequently omitted by these tools because of the heterogeneous information required for their application [36]. Overall, the inconsistencies that naturally rise in a "manual" variant interpretation process are inevitably reiterated. Automation of the interpretation process does not provide additional value when it is not based on precise and concrete definitions. This further motivates the effort described next.

3 Related Works

In the last years, several works have targeted specific domains all united by the lack of a solid and well-founded conceptual characterization of their characteristics. This was accomplished by proposing conceptual meta-models that provide general clarification and guidance on the understanding of the said domain. For instance, we report recent work in the context of fake news [51,6], Virtual Network Function Marketplaces [20], FAIR scientific datasets [8], or FAIR Digital Objects [47].

In the field of genomics, the use of conceptual models for specifying genomics-related processes has been explored. More specifically, conceptual modeling techniques have proven to be an effective tool to achieve high levels of concreteness and standardization. A recent work [7], has considered general genomic data types represented in datasets for analysis and connected them to an abstract conceptual representation, with the purpose to resolve their heterogeneity. Other modeling efforts have targeted the inherent temporal dimension associated with genomics data by mapping their evolution over time [15]; such an approach is particularly sensitive in cases of changes in variant interpretation due to gene-related data being updated [48]. Conceptual models have also been proposed to target other specific aspects in the use of multi-omics data for precision medicine [45] and for the identification of relevant and high-quality data records [38].

Aside from simple conceptual models, also ontological approaches have been attempted. Ferrandis et al. [32] promoted the use of foundational ontologies to avoid errors while creating and curating genomic domain models for personalized medicine. The approach of ontological clarification has been employed to support the explanation of complex domains such as human metabolic pathways [16] and the viral genome with the related events of infection, sampling, sequencing, and annotation for SARS-CoV-2 sequences [9]. Similarly, OntoRepliCov [27] showed an initial conceptual framework targeting the translation event during SARS-CoV-2 replication.

Despite the growing interest interest in conceptual models in the area of genomics and some technological efforts to gather and integrate different human variation data [31,12,54], to the best of our knowledge, the proposal presented here is the first explicit, reusable reference meta-model that targets the Variant Interpretation process.

4 A Meta-model for Variant Interpretation Guidelines

Let us begin by recalling that, as previously stated, variants can be classified according to a variety of interpretation guidelines. In the model proposed here (see Figure 1), each Guideline is defined by its *title*, *authors*, and its *applicability*, i.e., the specific context in which the guideline is applicable, such as for instance "Mendelian disorders" –a specific type of disorder–, or "copy number variants" –a specific type of variant. In addition, guidelines have a *URL* that points to the publication or file where they can be examined.

For the purpose of this model, we consider that the Variant only characteristic is an unique *identifier* (e.g., "rs556540177" for a single nucleotide variant, or "nsv3875336" for a copy number variant). This particualar domain oversimplification enabled us to reduce the complexity of the model. However, an extended version of this model that includes additional information such as the position of the Variant on the distinct genomic sequences, as well as the reference and alternative alleles exists. A ClassificationResult (e.g., "benign", "pathogenic", or "protective") is the classification obtained for a certain Variant using a specific Guideline.

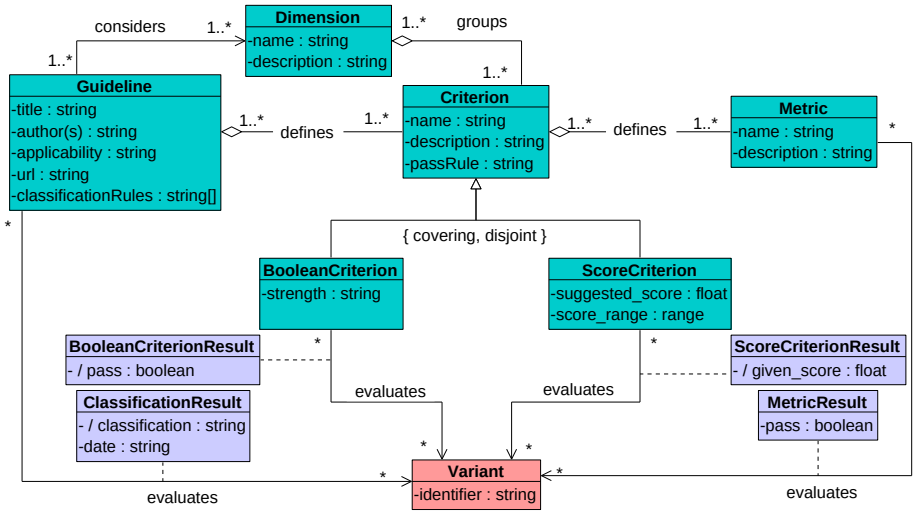

Fig. 1. Meta-model for variant interpretation guidelines. Concepts associated with the definition of clinical guidelines are depicted in green, the DNA-associated concepts (i.e., the variant that is interpreted by the clinical guideline) are depicted in red, and the concepts used to describe the results of interpreting a variant via a clinical guideline are depicted in lilac.

Each Guideline defines a number of criteria (i.e., a set of Criterion) to be evaluated in order to obtain the most adequate *ClassificationResult* for a Variant. This classification is calculated based on the *classificationRules* defined in a Guideline, which state the combination of criteria that must be met to achieve a specific *ClassificationResult*. An example of a classification rule is "pathogenic: PS1, benign: BP1", which specifies that a variant would be classified as "pathogenic" if the PS1 criterion was met, and as "benign" if the BP1 criterion was met.

A Criterion is decomposed into more specific aspects, called metrics. Each Metric evaluates to either a `true` or a `false` value for a particular Variant (i.e., representing a MetricResult). Similarly to Guidelines, each Criterion defines a specific rule, named *passRule*, which performs logical operations over the set of MetricResults to determine whether the Criterion is met. It is worth noting here that the same Metric can be used to calculate multiple criteria (as represented by the cardinalities between the Metric and the Criterion classes).

We recognize two different kinds of Criterion: the BooleanCriterion and the ScoreCriterion. The BooleanCriterion returns a `true` or `false` value (i.e., BooleanCriterionResult) and is defined by a *strength* that represents the extent to which the criterion supports a specific classification. For instance, a "strong" value indicates that the fulfillment of the criteria provides strong support for a certain classification, whereas a "moderate" value indicates

that the criteria only offers moderate support. The ScoreCriterion returns a numeric value (i.e., ScoreCriterionResult) and has a float *suggested_score* (e.g., "0.25") within a *score_range* (e.g., "[0, 0.45]").

Lastly, the Dimension groups distinct criteria that share given aspects. For instance, some criteria focus on evaluating specific characteristics of a Variant position in our DNA sequence, in which case we have a Dimension with the *name* "Variant position", and the *description* "Criteria that evaluate aspects of the variant location in the DNA". Making this common background explicit among various criteria improves interoperability among different Guidelines.

4.1 Example: PM1 criterion of the ACMG-AMP 2015 guidelines

Here, we provide an illustration of the use of the meta-model by instantiating the PM1 criterion of the ACMG-AMP 2015 variant interpretation guidelines (see Figure 2 for a textual description). This is one of 16 criteria that support the analysis of a variant's pathogenicity according to these guidelines. More specifically, it evaluates whether a variant is found in a region of our DNA known as a "mutational hotspot" (i.e., a DNA region that has a high frequency of pathogenic variants) and/or in a "protein domain" (i.e., a stable, independent part of a protein that can perform vital protein functions) that is critical for its correct functioning with no previously reported benign variations.

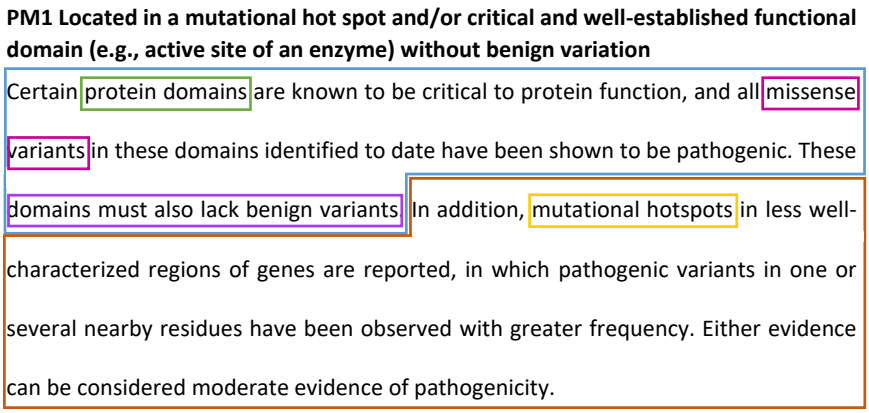

Fig. 2. The two criteria identified in the textual description of PM1 are highlighted in blue and brown frames. The four metrics identified in the textual description of PM1 are highlighted in green, pink, purple, and yellow frames. The blue-framed criterion comprises three metrics, whereas the brown-framed criterion only comprises one.

During the instantiation process of the PM1 criterion, two major issues emerged. The first one arises from the actual definition of the criterion, which appears to describe two distinct criteria rather than just one. Indeed, a variant

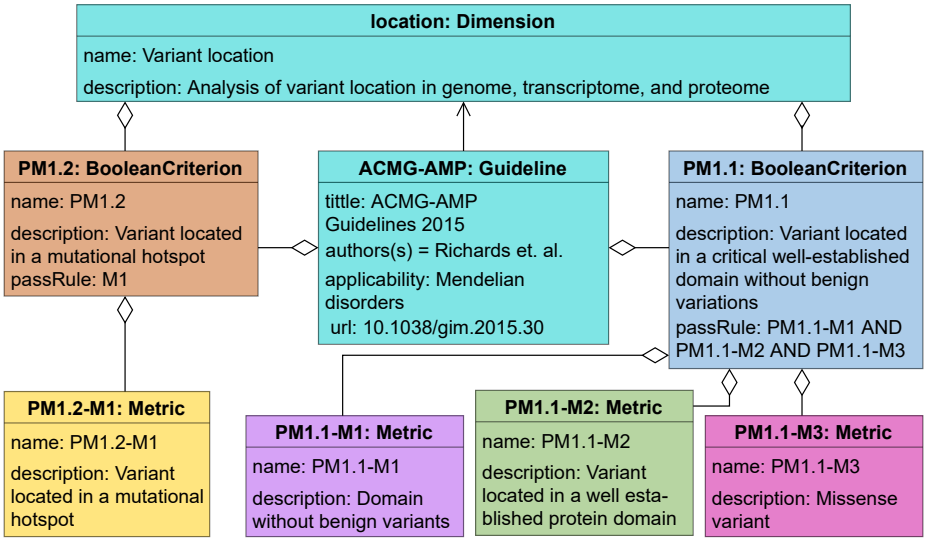

Fig. 3. The classes are depicted in the same colors used to highlight criteria and metrics in the textual description of PM1 in Fig. 2.

meets the PM1 criterion if it is discovered in a mutational hotspot, a functional domain without any known benign variations, or both of them. These scenarios provide different characteristics to be met and different requirements to be evaluated. Regardless of the fact that both hotspots and functional domains are genomic regions, they are of different types: hotspots are found in our DNA sequence, whereas domains are found in proteins. Furthermore, the absence of benign variation is only important for assessing the condition associated with protein domains. Therefore, the PM1 criterion descriptions collapse two different criteria. When they are considered separately, as promoted by our model, the evaluation of the criterion becomes clearer and simpler.

The second issue concerns the imprecise definition provided in PM1. According to this criterion, the variant must be "located in a mutational hot spot and/or critical and well-established functional domain without benign variations". However, a careful reading of the complete description of the criterion reveals that the part of PM1 regarding protein domains is only valid for missense variants (i.e., a variant that leads to an amino acid change in the protein sequence).

Figure 3 shows the resulting instantiation of the PM1 criterion. The criterion has been instantiated as two different Boolean criteria (PM1.1 and PM1.2) that evaluate to either `true` or `false`. On the one hand, if a missense variant (PM1.1-M3 metric) is located in a well-established functional domain (PM1.1-M2 metric) with no benign variants (PM1.1-M1 metric), the PM1.1 criterion evaluates to true –the *passRule* of the criterion is composed by the conjunction of these three metrics. On the other hand, if a variant is found in a DNA region known to be a mutational hotspot (PM1.2-M1 metric), the PM1.2 evaluates to true.

Here, the BooleanCriterion has a much simpler *passRule*, only including the measurement of M1.

The meta-model has allowed us to unpack and make explicit the constructs underlying the ACMG-AMP PM1 criterion, which were previously hidden in the convoluted nature of its textual description. This unpacking process was supported by the part-hood relationships defined between the Guideline and Criterion classes, and between the Criterion and Metric classes. These part-hood relationships are made explicit via the formulas defined in the *passRule* and *ClassificationRules* attributes. A Criterion's classification result is based on the evaluation of its metrics. Similarly, a Guideline's classification result is based on the evaluation of its criteria. Our metamodel enables the decomposition of variant interpretation guidelines into more precise constructs, which can serve as a solid foundation for clinical guidelines operationalization.

5 Variant Misinterpretation Patterns

The meta-model characterizes the constructs and underlying structure of interpretation guidelines. This characterization has led us to the identification of four patterns that hinder the variant (mis)interpretation process. These patterns highlight the main inconsistencies in the interpretation processes when used by different experts; they also elucidate the disparities in the variant classification results. We have identified four different patterns: 1) the use of a single Metric leading to different MetricResults; 2) the use of a single Criterion measured according to different Metrics; 3) the use of a Guideline with diverse Criteria; and 4) the use of one Criteria with different purposes within diverse Guidelines. All such patterns are allowed in the meta-model and represented by several real-world examples; however, they are at the basis of situations unclear/incoherent interpretations of variants. Details and examples are provided in the next sections.

5.1 Same metric – different metric results

The lack of data sharing is a significant issue in genomics [39]. Indeed, differential access to privately stored data is one of the most common causes of discrepancies in variant interpretation [13,21]. Because of this, different experts may evaluate the same criterion's metric differently depending on the data they have access to.

Let us consider the following example. Determining whether a variant co-occurs with a pathogenic variant is frequently regarded as proof of the benignity of the variant under investigation [43]. It is common for laboratories that perform genetic testing to have their own variant repository that they do not share publicly [33]. As a result, one laboratory may have identified cases in which the variant co-occurs with a pathogenic variant while another laboratory may not hold this information [13]. Consequently, when the metric "*The variant co-occurs with a pathogenic variant(s)*" is evaluated, different metric results may be obtained, depending on the data that the laboratory uses.

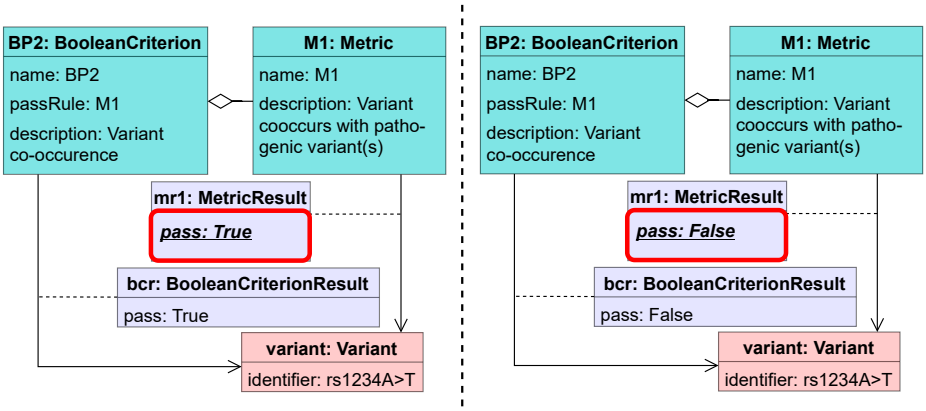

Fig. 4. Example model of pattern "Same metric – different metric result"

Different metric results will influence whether or not a particular criterion is met. A practical example of this situation is depicted in Figure 4. The variant rs1234A>T has met the criterion that evaluates variant co-occurrence (BP2 criterion) in the first scenario because the metric "*The variant co-occurs with pathogenic variant(s)*" (metric M1) has been met. However, in the second scenario, the variant fails the BP2 criterion because the metric M1 is not met, thereby impacting also the BooleanCriterionResult pass value. Our meta-model has allowed us to identify that the misinterpretation of the BP2 criterion is due to different metric evaluation results.

5.2 Same criterion – different metrics

Interpretation guidelines have contributed to the standardization of the variant interpretation process. However, due to the lack of specificity in these guidelines, different experts may apply the same criterion differently [22]. This indicates, according to our meta-model, that different metrics have been employed to evaluate the same guideline's criterion.

This is especially common when determining a variant's allele frequency [13]. Variant interpretation guidelines frequently recommend using the allele frequency of the variant as a benignity criterion if it is *greater than expected for that specific disorder*. Such a definition makes the frequency's cutoff entirely dependent on the knowledge and experience of the expert performing the interpretation [25]. As a result, given the criterion for evaluating allele frequency, one expert could define a metric that states, for instance, that "*the variant should have an allele frequency greater than 0.5%*", whereas an alternative expert – with a stricter approach – could define a different metric stating that "*the variant should have an allele frequency greater than 1%*". This difference in metrics may obviously result in different assessments of whether or not the same criterion is met.

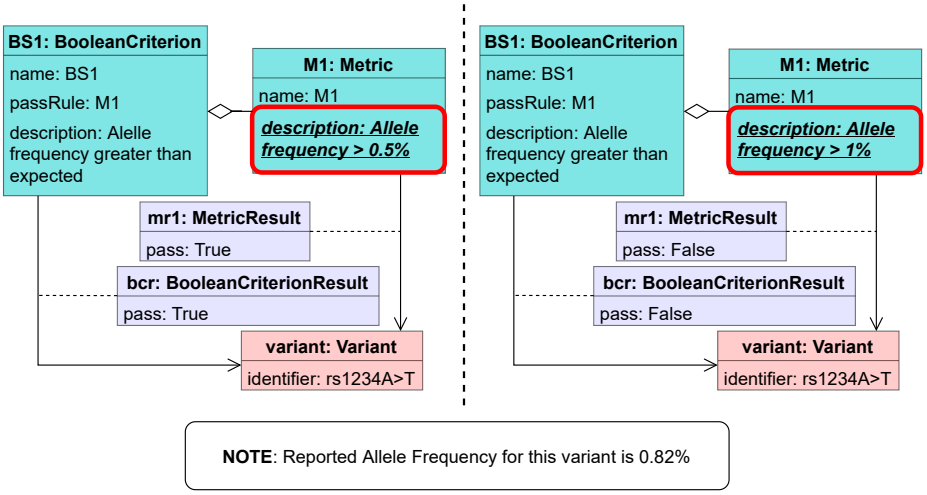

Fig. 5. Example model of pattern "Same criterion – different metrics"

Figure 5 depicts an actual instance model of this situation. When the criterion BS1 (evaluating whether "*the allele frequency of the variant is greater than expected for the disorder*") is applied to the rs1234A>T variation, it produces different results, depending on the different definitions of the (only) metric which this criterion depends on. Again, our meta-model is able to pinpoint clearly the origin of criterion assessment differences.

5.3 Same guideline – different criteria

Most common misinterpretations occur when merging results from different sources that follow different guidelines. One would expect that this could not happen *within* the context of a specific guideline, as these intend to create a well-defined framework for selecting the most appropriate interpretation for a variation. Surprisingly, differences in interpretation results are common even when using the same interpretation guideline [3,4]. This is related to the fact that laboratories that perform the "interpretation" activity may be unable (for diverse reasons–economic, time-related, or motivational) to apply all of the criteria specified in the guidelines.

This is frequently the case in functional studies. Many variant interpretation guidelines recommend using well-conducted functional studies to assess the potential impact of a variation in a gene or gene product [43,11]. This type of research, however, is extremely difficult to pursue due to the significant monetary and time investment required. As a result, only 36% of clinical experts apply this criterion during the variant interpretation process [55].

Because functional studies provide strong evidence of the pathogenicity of the variant, the choice of the expert to use this type of evidence will have a significant impact on the interpretation of the variation. This is especially important for

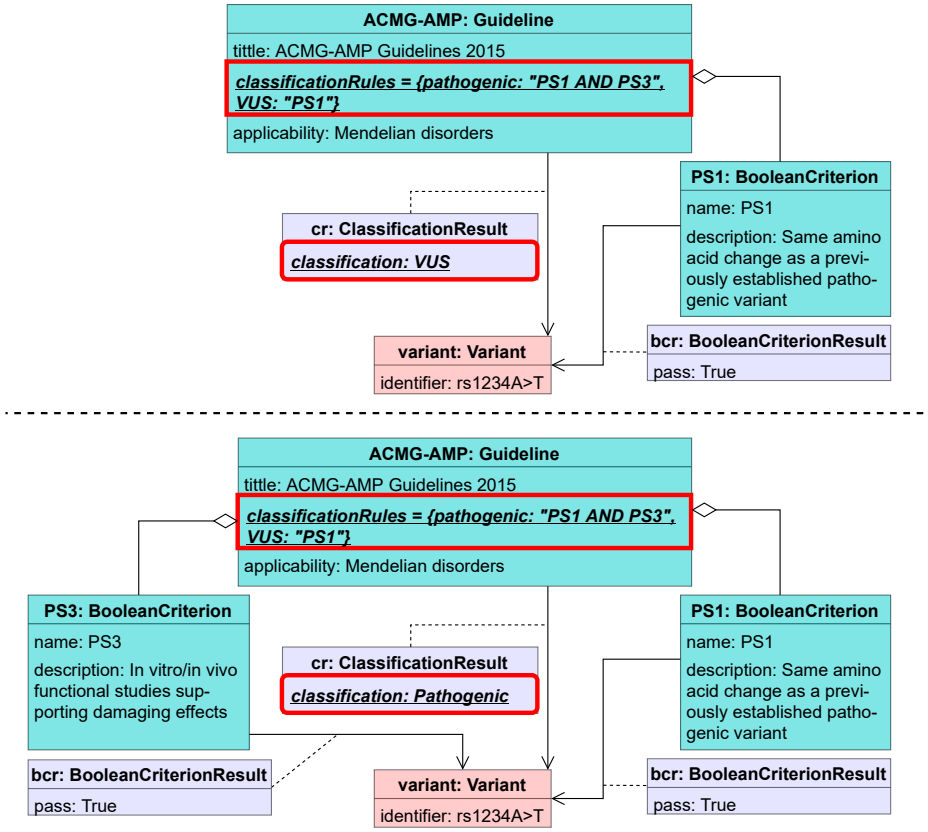

Fig. 6. Example model of pattern "Same guideline – different criteria"

variants whose significance is unclear, and a functional study can determine whether the variant should be discarded as benign or further investigated for its potential to cause disorder [18,5].

The impact that the used criteria can have on the interpretation of a variation is demonstrated practically in Figure 6. The expert in the top scenario only considered criterion PS1, thus concluding that the variant has an Uncertain Significance (VUS) based on that information. However, the expert in the bottom scenario considered both PS1 and PS3; according to the classification rule that assigns the "pathogenic" value when both PS1 and PS3 hold, or the "VUS" value when only PS1 holds, this expert concluded that the variant should be classified as pathogenic. The additional evidence provided by functional studies (criterion PS3) was fundamental in this case. The meta-model provides a clear representation of each expert's interpretation process and pinpoints the source of inconsistencies in the interpretation of variant rs1234A>T.

5.4 Same criterion – different guidelines

Different variant interpretation guidelines establish different criteria and metrics depending on their applicability. Nevertheless, there are well-established criteria that usually appear in multiple guidelines.

In clinical guidelines, each criterion is defined using two alternative approaches: Boolean-based or score-based. When a Boolean criterion is used, the criterion is either met or not. When a score criterion is used, instead, a criterion is accepted if its associated value falls within a predefined range. Consequently, even when guidelines include the same criterion, its assessment may be different depending on the approach adopted by the guideline.

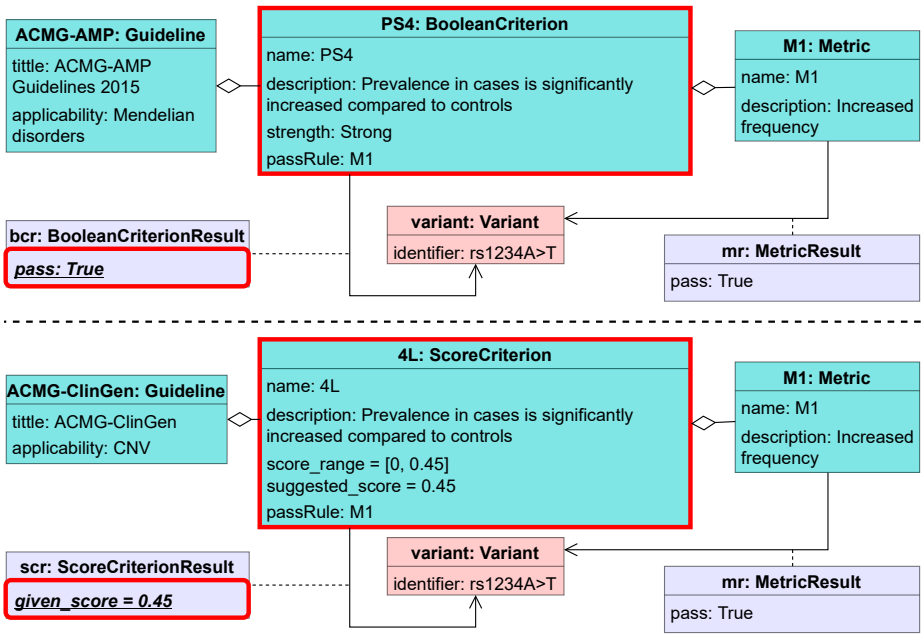

Fig. 7. Example model of pattern "Same criterion – different guidelines"

A typical case when this difference emerges involves the criterion that evaluates whether a variant is more frequent in cases than in controls. The criterion is evaluated by the ACMG-AMP 2015 guidelines as a Boolean criterion [43], and in the ACMG-ClinGen as a score criterion [44]. Figure 7 illustrates the example. In the ACMG-AMP Guideline, the criterion PS4 analyzes whether the frequency of the variant rs1234A>T is increased in affected individuals – by means of the metric M1. The M1 result evaluates as true and, consequently, the PS4 criterion results are also evaluated as true. In the ACMG-ClinGen Guideline, the equivalent criterion 4L evaluates the same metric for the same variant. In this case,

the criterion result is a particular score (0.45), whose value is obtained based on the score range and the suggested score stated in the criterion definition.

Our meta-model clearly illustrates the differences between both guidelines and - in general - allows experts to identify variant interpretation differences that arise from the use of different approaches for variant interpretation.

6 Discussion

Variant interpretation is a critical step in achieving better diagnoses and treatments based on each individual's genomic information. However, the imprecise and vague nature of the variant interpretation guidelines poses difficulties in its application in a real clinical setting. We have used a conceptual modeling approach to define a meta-model that allows us to identify the structure and constructs behind interpretation guidelines.

With the proposed meta-model, we have defined and explained the common framework for representing various types of guidelines (Section 4); we then identified patterns of misinterpretation of variants (Section 5); finally, the previous results enabled us to disentangle intricate details of existing clinical guidelines, as we analyzed in the examples of the previous section. Below, we summarize the lessons learned during this process.

Unpacking variant classification results: Differences in variant interpretation can have important consequences on a patient's health. The reason behind these differences sometimes is not the use of a different guideline or criterion but a conflicting evaluation of the same criterion. Thanks to the description of a criterion as an aggregation of metrics, we are not only able to identify a different evaluation of a criterion but the specific metric that has caused such a difference. Section 5.1 illustrates this case. This allows for a precise unpacking of the variant classification results.

Disambiguating criterion definitions: Because the interpretation guidelines are often not clear enough for their unambiguous application, various experts will use different measurements to determine whether a criterion is met. As seen in Section 5.2, the metric definition has allowed us to identify the collection of constructs an expert uses to assess a certain criterion. This enables us to provide a standard framework for comparing various interpretations of the same criterion.

Clarifying interpretation guidelines application: A precise set of criteria are specified in the interpretation guidelines to direct the classification outcome. As Section 5.3 shows, not all experts employ all criteria, which makes it difficult to derive the precise procedure that was used. The meta-model enables a precise characterization of the particular criteria applied for variant interpretation as well as the components assessed in each criterion, enabling full traceability of the outcomes.

Making connections explicit: The 22 interpretation guidelines currently available have important differences in their applicability, the criteria considered most important to assess the role of a variant in the disorder process, or even in their

approach for evaluating such criteria (boolean or score). Precisely identifying the differences and commonalities among the guidelines is key to comparing the interpretation approach followed by different experts and the possible implications for the classification results. Section 5.4 reflects how the meta-model has allowed us to make explicit connections among different interpretation guidelines.

Operationalization of guidelines: Clinical guidelines were originally defined in an abstract manner thus hampering their direct operationalization. The generated conceptual schema poses the basis for building workflows that systematically: 1) explain the complex interpretation domain (on the lines of [19]) and the related process in place (a sort of process explainability [1]); 2) highlight current differences, inconsistencies, and misinterpretations; 3) propose refinements to current criteria and metrics; and 4) derive a complete operationalization of the guidelines' application process. A conceptual model can serve as the foundation for operationalizing clinical guidelines by making them more accessible, guiding decision-making, facilitating interdisciplinary collaboration, and encouraging continuous improvement. As a result, inconsistencies in their application will be reduced.

Current limitations of the meta-model: During the development and use of the meta-model, we identified four limitations. First, there are external elements that may have an impact on classification results, but they are not represented. These elements include, for example, the fact that some variants are pathogenic only when appear in combination with other variants, or that other variants may overcome a variant's pathogenic effect. Second, variant interpretation has not been examined in the context of complex disorders in this work. In these disorders, the existence of many variants are required to cause the manifested disorder. Extra factors such as penetrance and population specificities must be considered also, but they are not included in this first version of the model. Third, the actors participating in the interpreting process are not modeled. Knowing who performed the interpretation, what annotation tool was used, or what information they relied on to evaluate each criterion helps increasing the interoperability and reproducibility of the interpretation results. Fourth, we represent the *classificationRules* as an array of strings. Although this approach works correctly, we are aware that more appropriate appropriate approaches exist. For instance, specific classes that better capture the nature of these rules.

7 Conclusion and Future Outlook

In this paper, we proposed a novel meta-model for the representation of the DNA Variant Interpretation Guidelines. Variant interpretation is a very common process in the working routine of clinicians and geneticists and it is of critical importance that it is managed in a correct way to ensure patients well-being. Unfortunately, current practice still presents many shortcomings; the presence of several guidelines with diverse criteria and metrics – possibly based on different approaches or with apparent discrepancies – is hampering the reliability of the interpretation results.

Paving the way to a complete standardization and systematization of this process, here we proposed a meta-model that aims to explain and clarify the morphology of interpretation guidelines and their internal elements. Additionally, we proposed a set of patterns in which these guidelines led to the potential misinterpretation of variants. These patterns reveal common challenges encountered when interpreting variants and each of them is associated with a practical use case where the pattern arose. Finally, we discussed lessons learned during the modeling effort and how these reflect on the presented problematic use cases.

In the future, we plan to address the meta-model's limitations identified above. First, we intend to represent the variant's genomics context to show how the existence of other variants may influence the variant's classification. Second, our model will incorporate a classification of variant groups that operate together to produce a disorder. This will facilitate the interpretation of complex disorders. Third, we will incorporate a detailed description of all the steps that precede variant interpretation as well as the actors that conducted the interpretation process. Finally, a new entity capturing the complexity and interconnections of the classification rules will be defined.

In addition, we plan to thoroughly expand the patterns catalog, proposing operational rules to avoid such incorrect situations to occur. As previously discussed, this preliminary meta-model effort will be applied to practical frameworks for two main purposes. First, we aim to explain the complex variant interpretation process, reporting differences, inconsistencies, and misinterpretations. Second, we aim to propose refinements to current criteria and metrics and completely operationalize the guidelines' application process.

Acknowledgements This work was supported by the Valencian Innovation Agency and Innovation through the OGMIOS project (INNEST/2021/57), the Generalitat Valenciana through the CoMoDiD project (CIPROM/2021/023) and ACIF/2021/117, and the Spanish State Research Agency through the DELFOS (PDC2021-121243-I00,MICIN/AEI/10.13039/501 100011033) and SREC (PID 2021-123824OB-I00) projects, and co-financed with ERDF and the European Union Next Generation EU/PRTR.

References

1. Adadi, A., et al.: Peeking inside the black-box: A survey on explainable artificial intelligence (xai). IEEE Access **6**, 52138–52160 (2018)
2. Agaoglu, N., et al.: Consistency of variant interpretations among bioinformaticians and clinical geneticists in hereditary cancer panels. European Journal of Human Genetics **30** (02 2022)
3. Amendola, L.M., et al.: Performance of ACMG-AMP Variant-Interpretation Guidelines among Nine Laboratories in the Clinical Sequencing Exploratory Research Consortium. The American Journal of Human Genetics **98** (05 2016)
4. Amendola, L.M., et al.: Variant classification concordance using the ACMG-AMP variant interpretation guidelines across nine genomic implementation research studies. The American Journal of Human Genetics **107**(5), 932–941 (2020)

5. Anderson, C., et al.: How functional genomics can keep pace with vus identification. Frontiers in Cardiovascular Medicine **9** (Jul 2022)
6. Belloir, N., et al.: Characterizing fake news: A conceptual modeling-based approach. In: Conceptual Modeling: 41st International Conference, ER 2022, Hyderabad, India, October 17–20, 2022, Proceedings. pp. 115–129. Springer (2022)
7. Bernasconi, A., et al.: A comprehensive approach for the conceptual modeling of genomic data. In: Ralyté, J., et al. (eds.) Conceptual Modeling. pp. 194–208. Lecture Notes in Computer Science, Springer International Publishing (2022)
8. Bernasconi, A., et al.: Ontological representation of fair principles: A blueprint for fairer data sources. In: Proceedings of the 35th International Conference on Advanced Information Systems Engineering (CAiSE 2023) (2023)
9. Bernasconi, A., et al.: Semantic interoperability: ontological unpacking of a viral conceptual model. BMC Bioinformatics **23**(Suppl 11), 491 (2022b)
10. Booch, G., et al.: The unified modeling language. Unix Review **14**(13), 5 (1996)
11. Brnich, S., et al.: Recommendations for application of the functional evidence PS3/BS3 criterion using the ACMG/AMP sequence variant interpretation framework. Genome medicine **12**, 3 (12 2019)
12. Canakoglu, A., et al.: Genosurf: metadata driven semantic search system for integrated genomic datasets. Database **2019** (2019)
13. Furqan, A., et al.: Care in specialized centers and data sharing increase agreement in hypertrophic cardiomyopathy genetic test interpretation. Circulation: Cardiovascular Genetics **10**(5), e001700 (2017)
14. Gao, P., et al.: Challenges of providing concordant interpretation of somatic variants in non-small cell lung cancer: A multicenter study. J Cancer **10**(8), 1814–1824 (2019)
15. García S., A., et al.: The challenge of managing the evolution of genomics data over time: a conceptual model-based approach. BMC Bioinformatics **23**(11), 472 (2022)
16. García S, A., et al.: An ontological characterization of a conceptual model of the human genome. In: De Weerdt, J., et al. (eds.) Intelligent Information Systems. pp. 27–35. Springer International Publishing, Cham (2022)
17. Genetics, A.: Scheme for autosomal dominant and x-linked mendelian diseases, Ambry Genetics (2015). `https://submit.ncbi.nlm.nih.gov/ft/byid/zfkfvckw/mid-7377_ambry_classification_scheme_oct_2015.pdf`, accessed: 2023-05-24
18. Guidugli, L., et al.: Functional assays for analysis of variants of uncertain significance in brca2. Human Mutation **35**(2), 151–164 (2014)
19. Guizzardi, G., et al.: Ontological unpacking as explanation: the case of the viral conceptual model. In: Conceptual Modeling: 40th International Conference, ER 2021, Virtual Event, October 18–21, 2021, Proceedings 40. pp. 356–366. Springer (2021)
20. Guizzardi, R., et al.: A reference conceptual model for virtual network function online marketplaces. In: Conceptual Modeling: 38th International Conference, ER 2019, Salvador, Brazil, November 4–7, 2019, Proceedings 38. pp. 302–310. Springer (2019)
21. Harrison, S.M., et al.: Scaling resolution of variant classification differences in ClinVar between 41 clinical laboratories through an outlier approach. Human Mutation **39**(11), 1641–1649 (2018)
22. Harrison, S.M., et al.: Clinical laboratories collaborate to resolve differences in variant interpretations submitted to ClinVar. Genetics in Medicine **19**(10), 1096–1104 (2017)

23. Karczewski, K.J., et al.: The mutational constraint spectrum quantified from variation in 141,456 humans. https://doi.org/10.1101/531210
24. Kim, Y.E., et al.: Challenges and considerations in sequence variant interpretation for mendelian disorders. Annals of Laboratory Medicine **39**, 421 (09 2019)
25. Kim, Y.E., et al.: Challenges and considerations in sequence variant interpretation for mendelian disorders. Annals of Laboratory Medicine **39**, 421 (09 2019)
26. Kopanos, C., et al.: VarSome: the human genomic variant search engine. Bioinformatics **35**(11), 1978–1980 (10 2018)
27. Laddada, W., et al.: OntoRepliCov: an Ontology-Based Approach for Modeling the SARS-CoV-2 Replication Process. Procedia Computer Science **192**, 487–496 (2021)
28. Lebo, M.S., et al.: Data sharing as a national quality improvement program: reporting on brca1 and brca2 variant-interpretation comparisons through the canadian open genetics repository (cogr). Genetics in Medicine **20**(3), 294–302 (2018)
29. Li, Q., et al.: InterVar: Clinical Interpretation of Genetic Variants by the 2015 ACMG-AMP Guidelines. The American Journal of Human Genetics **100**(2), 267–280 (2017)
30. Luo, X., et al.: ClinGen myeloid malignancy variant curation expert panel recommendations for germline RUNX1 variants. Blood Advances **3**(20), 2962–2979 (2019)
31. MacDonald, J.R., et al.: The database of genomic variants: a curated collection of structural variation in the human genome. Nucleic Acids Research **42**(D1), D986–D992 (2014)
32. Martínez Ferrandis, A.M., et al.: Applying the principles of an ontology-based approach to a conceptual schema of human genome. In: Ng, W., et al. (eds.) Conceptual Modeling. pp. 471–478. Springer Berlin Heidelberg, Berlin, Heidelberg (2013)
33. Mighton, C., et al.: Data sharing to improve concordance in variant interpretation across laboratories: results from the canadian open genetics repository. Journal of Medical Genetics **59**(6), 571–578 (2022)
34. Naithani, N., et al.: Precision medicine: Uses and challenges. Medical Journal Armed Forces India **77** (07 2021). https://doi.org/10.1016/j.mjafi.2021.06.020
35. Nicora, G., et al.: CardioVAI: An automatic implementation of ACMG-AMP variant interpretation guidelines in the diagnosis of cardiovascular diseases. Human Mutation **39** (10 2018)
36. Nicora, G., et al.: A machine learning approach based on ACMG/AMP guidelines for genomic variant classification and prioritization. Sci Rep **12**(1), 2517 (2022)
37. Niehaus, A., et al.: A survey assessing adoption of the ACMG-AMP guidelines for interpreting sequence variants and identification of areas for continued improvement. Genet Med **21**(8), 1699–1701 (2019)
38. Palacio, A.L., et al.: A method to identify relevant genome data: Conceptual modeling for the medicine of precision. In: Trujillo, J.C., et al. (eds.) Conceptual Modeling. pp. 597–609. Lecture Notes in Computer Science, Springer International Publishing (2018)
39. Powell, K.: The broken promise that undermines human genome research. Nature **590**, 198–201 (02 2021)
40. Ramdaney, A., et al.: Beware the laboratory report: Discrepancy in variant classification on reproductive carrier screening. Genetics in Medicine **20**, 374–375 (10 2017)
41. Reis-Filho, J.S.: Next-generation sequencing. Breast Cancer Research **11**(3), S12 (Dec 2009)

42. Reyes Román, J.F., et al.: Applying conceptual modeling to better understand the human genome. In: Comyn-Wattiau, I., et al. (eds.) Conceptual Modeling. pp. 404–412. Lecture Notes in Computer Science, Springer International Publishing (2016)
43. Richards, S., et al.: Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the american college of medical genetics and genomics and the association for molecular pathology. Genetics in medicine **17**(5), 405–423 (2015)
44. Riggs, E., et al.: Technical standards for the interpretation and reporting of constitutional copy-number variants: a joint consensus recommendation of the american college of medical genetics and genomics (acmg) and the clinical genome resource (clingen). Genetics in Medicine **22**, 1–13 (11 2019)
45. S., A.G., et al.: A conceptual model-based approach to improve the representation and management of omics data in precision medicine. IEEE Access **9**, 154071–154085 (2021)
46. Scott, A.D., et al.: CharGer: clinical characterization of germline variants. Bioinformatics **35**(5), 865–867 (2019)
47. Bonino da Silva Santos, L.O., et al.: Towards a conceptual model for the fair digital object framework. In: Proceedings of the 13th International Conference on Formal Ontology in Information Systems (FOIS 2023) (2023)
48. So, M.K., et al.: Reinterpretation of BRCA1 and BRCA2 variants of uncertain significance in patients with hereditary breast/ovarian cancer using the ACMG/AMP 2015 guidelines. Breast Cancer **26**, 510–519 (2019)
49. Tavtigian, S.V., et al.: Modeling the ACMG/AMP variant classification guidelines as a bayesian classification framework. Genetics in Medicine **20**(9), 1054–1060 (2018)
50. Vihinen, M.: Problems in variation interpretation guidelines and in their implementation in computational tools. Molecular Genetics & Genomic Medicine **8**(9), e1206 (2020)
51. Weiss, A.P., et al.: Toward a comprehensive model of fake news: A new approach to examine the creation and sharing of false information. Societies **11**(3), 82 (2021)
52. Whiffin, N., et al.: Cardioclassifier: disease- and gene-specific computational decision support for clinical genome interpretation. Genetics in Medicine **20**(10), 1246–1254 (2018)
53. Zeggini, E., et al.: Translational genomics and precision medicine: Moving from the lab to the clinic. Science **365**(6460), 1409–1413 (2019)
54. Zhang, J., et al.: The international cancer genome consortium data portal. Nature biotechnology **37**(4), 367–369 (2019)
55. Zirkelbach, E., et al.: Managing variant interpretation discrepancies in hereditary cancer: Clinical practice, concerns, and desired resources. Journal of Genetic Counseling **27**(4), 761–769 (2018)