CONFIDENTIAL

ACMG Journals Technical Reviewer Guidelines

VERSION 1.0.0
PETER FREEMAN
LISA LANSDON

APPROVED BY
ROBERT STEINER
DAVID MILLER

CONTRIBUTIONS FROM
BINU PORATH
MARI MORI

Acknowledgements

Heidi Rehm - ACMG WG; Genome Medical, Cell Genomics

Commented [2]: If you common on or edit this manual, and are not a mamber of the GIM technical team, please add your name, the working groups(s) you represent and any journals you represent.

Access to the document

https://github.com/openvar/GIM_Technical_Editing

Technical Editor Checklist

1. For manuscripts that contain any gene information (e.g. gene symbols), or genetic variation information.
    a. A genome build must be specified.
        i. GRCh38 is preferred however some labs still use GRCh37.
        ii. hg19 and hg38 can also be stated.
        iii. hg19 is not equivalent to GRCh37. The mitochondrial chromosome is different. NC_012920.1 is used by GRCh37, GRCh38 and hg38. NC_001807.4 is used by hg19.
        iv. If mitochondrial variants are included, ensure that the correct reference sequence is stated.
    b. Ideally, the genome build source should also be specified
        i. Ideally a genbank ID e.g. GCA_000001405.29 or a RefSeq ID e.g. GCF_000001405.40 for the genome build should be stated, or web link to the genome build that was used in an analysis should be provided e.g. https://console.cloud.google.com/storage/browser/genomics-public-data/resources/broad/hg38/v0/

2. Gene Information (see Guidance on gene names, symbols, and IDs below)
    a. All genes must be represented by a gene symbol e.g. *COL1A1* and the symbol must be italicised (Note: Use of the term "gene name" applied to a gene symbol is incorrect. For example the *COL1A1* is a gene symbol and "collagen type I alpha 1 chain" is the corresponding gene name).
    b. Gene symbols must be accompanied by a stable HGNC gene ID, e.g. *COL1A1* (HGNC:2197). Other stable gene IDs are available for example: MIM ID; genbank ID; Ensembl ID, however, these are intrinsically linked in the HGNC gene ID record and we insist on the HGNC ID. The IDs can be found at https://www.genenames.org/.
    c. If variants are described at the transcript level (i.e. c. or r.) a reference sequence must also be provided with the gene symbol and ID (see Associating genes with transcript reference sequence IDs).
    d. Does the gene have a reported gene-disease relationship in a publication and/or database such as ClinGen, OMIM or GenCC?
        - If not, ask authors to follow the steps in Strande *et al*., paper (PMID: 28552198) to determine the gene-disease association or update/review an existing claim.
        - Authors should reference applicable manuscripts.

2. Protein variant descriptions
    a. Check that protein variants (p.) are not described without a corresponding c. or r. description somewhere in the manuscript.

b. Ensure that p. descriptions inferred from c. descriptions have parentheses (see “When to use parentheses”).

3. Transcript variant descriptions
    a. Ensure a reference sequence is provided.
    b. If the transcript reference is the same as a MANE Select or MANE Plus Clinical reference for the gene, but is an older version, encourage authors to update their description to the appropriate MANE transcript where possible.
        - If an alternative transcript to MANE Select or MANE Plus Clinical is preferred by authors, the rationale for electing to use said transcript should be stated in the manuscript.
    c. Ensure the descriptions are correctly formatted (see “Common nomenclature errors”).

4. Required variant descriptions
    a. Authors should use and provide the version of the HGVS and ISCN nomenclature that they used. The HGVS version can be found at https://hgvs-nomenclature.org/stable/ and the ISCN version at https://iscn.karger.com/
    b. Novel variants identified through genome sequencing (including transcriptome sequencing) should be fully described with a genomic (g.), transcript (c.) and protein (p.) descriptions (unless intergenic or spanning more than the boundaries of all transcripts for a single gene).
    c. Every document, including supplementary documents, must be self-contained (i.e. every variant stated in an individual document must be described at the g. c. and p. at least once in that document and must not rely on additional documents to provide any of these descriptions).
    d. Use of NM_# with version number is required.
        - Specify at least once for each gene that has a corresponding variant in the manuscript.
        - Specify with each variant only if authors are discussing more than one transcript.
        - Note: MANE select / MANE Plus Clinical is preferred.]
    e. Use of NC_# and genomic coordinates are required.
        - Although not required in the body of the manuscript, the NC_#. must be specified in a table or supplemental table.
    f. Variants provided in a table must be findable using text searching (i.e. tables including variants must be text based and not an image).
    g. If the sequence of an mRNA is known, an r. description can be included as well.
    h. If the data are 3rd party (i.e. taken from another manuscript or database entry), the authors must still use valid HGVS. Also, if the 3rd party includes a g. description (e.g. ClinVar) then authors should still provide it. If not, then VariantValidator can be used in many cases to generate a g. description which should then be included. The source of the descriptions must be provided, i.e. a reference associated with the reproduced variant description, or a database ID.

Commented [3]: Under review pending feedback from the ACMG committee

Commented [4]: I think we may consider releasing this as a V1. Noting that this is considered for review in early 2026 and then release an update to the TE manual in 2026. This will stop slowing us down :)

i. Sequence-level variant nomenclature must be HGVS compliant, and larger deletions and duplications list both the ISCN-like and HGVS-like nomenclature.
j. The size (i.e., bp, kb, or Mb) of each CNV should be specified as well as clinically relevant genes within the deleted or duplicated interval.
k. For manuscripts reporting a previously reported variant or variants in a non-HGVS format, the variant descriptions must be updated into the HGVS format and that source of the report cited, and a note of the historic variant description(s) is (are) provided.
l. For pharmacogenetic testing, appropriate star allele nomenclature and the corresponding PharmVar ID (*i.e.,* PVID; see https://www.pharmvar.org/) must be provided. Authors may also include the variants which comprise the star allele haplotype in accordance with HGVS nomenclature guidelines, however, doing so is only required if the manuscript discusses the specific function of one or more of the specific variants which comprise the star allele or are describing the identification of a new star allele.

5. External annotations (optional addition by authors, not required) do not negate the requirement for HGVS descriptions (i.e. these are annotations not variant descriptions).
    a. Authors should be encouraged to add external annotation sources to their variants where relevant. These can include:
        - dbSNP ID
        - ClinGen Allele registry ID (which is a recommended annotation)
        - ClinVar or LOVD ID (one or both of these is a recommended annotation)
    b. Authors may not use an external annotation in place of the correct HGVS. The annotations must complement the HGVS (and/or ISCN nomenclature).
    c. If an annotation is provided, check that the annotation matches the associated HGVS description.
    d. Where such annotations are provided, the tool should be correctly cited.
    e. **Use of Sequence Ontology (SO) terms is encouraged where appropriate -** SO is a structured, controlled vocabulary designed to describe the features of biological sequences and the impact of genetic variation on those features. SO provides standardized terms that categorize how sequence changes affect specific genomic elements like coding regions, introns, exons, splice sites, and regulatory domains. We encourage authors to apply SO terms to convey the consequences of variation in a consistent and biologically meaningful way when in circumstances where essential information cannot be extrapolated from a variant description. For example, when referring to diseases associated with imprinting, an HGVS or ISCN variant description should be accompanied by a SO term and ID such as paternally_imprinted (SO:0000136)

Commented [5]: what defines "larger". Do you want to say variants large enough to be observed on a karyotype?

Commented [6]: Good point. We need to coordinate with the ACMG document here. Is it defined in the manuscript yet???

Commented [7]: I like Heidi's suggestion. We can either incorporate it into this version of our guidelines as a simplified description of when to use ISCN, or we can expand the description by slightly modify the phrasing that is used by ISCN 2024:

In ISCN 2024: "The nomenclature should be as precise as possible for the technologies and bioinformatic algorithms used. ISCN and HGVS nomenclature can be used together to describe a structural or copy number variant".

Proposed modification: "The nomenclature should be as precise as possible for the technologies and bioinformatics algorithms used. Sequence-level variant nomenclature must be HGVS-compliant, and multi-nucleotide alterations (e.g., copy number variants, structural variants, etc.) can be described by both ISCN and HGVS nomenclature as appropriate."

or maternally_imprinted (SO:0000135) to accurately define the origin of the imprinted region. Other use cases include, but are not limited to, short_tandem_repeat_contraction [SO:0002163] short_tandem_repeat_expansion [SO:0002162], cryptic_splice_acceptor (SO:0001570) and cryptic_splice_donor [SO:0001571].

6. Validation reports should be included with all manuscripts. Technical editors request these if they are not included. For up to 10 variants, the file can be VariantValidator pdf files. When > 10 variants are included (or alternatively), the validation file must be a VariantValidator batch report.
    a. Editors should check that the validation file is submitted.
    b. Editors should check that the variants stated in the text match the report.
    c. The report can guide the editor with respect to nomenclature issues in the text.
    d. For large numbers of variants, if a discrepancy is seen, the author is responsible for ensuring the descriptions in the text match the validation files and ensure errors in the body text identified by the TEs are fixed.

Commented [8]: Did we want to amend this to VariantValidator only?

Commented [9]: Yes

7. Have the variants been classified in accordance with current standards and guidelines?
    a. Authors should accurately cite the classification guidelines which they have utilized.
    b. For research reporting classification of variants or reclassification of variants, we require authors to evaluate the variants using the ACMG/AMP classification system for small variants (PMID: 25741868) and the ACMG/ClinGen classification system for CNVs (PMID: 31690835). ***Authors should classify variants in this circumstance as if the gene-disease relationship is already established and clarify that this approach was taken in the text. Manuscripts that are working to establish a novel gene-disease relationship should evaluate the strength of a gene-disease relationship based on the Strande et al., (PMID: 28552198) framework and include supporting documentation within their manuscript.***
    c. In addition to the classification, the variant scoring results (e.g., "PM2", or " 2A = 1.0" in the case of CNV classifications) should be listed in a table (this can be a supplementary table, and should cross reference the variant table, as described in the author guidelines https://www.sciencedirect.com/journal/genetics-in-medicine/publish/guide-for-authors)

8. Final formatting notes to Technical Editors
    a. Authors should avoid unnecessary punctuation of their variant descriptions.
    b. The following examples (or any similar examples) are incorrect and should instead be written as NM_000088.4:c.589G>T p.(Gly197Cys) (Note, simply separate with whitespace)
        - NM_000088.4:c.589G>T;p.(Gly197Cys)

- NM_000088.4:c.589G>T, p.(Gly197Cys)
- NM_000088.4:c.589G>T:p.(Gly197Cys)

c. An exception is that it is appropriate for a period to follow the final description at the end of a sentence e.g. NM_000088.4:c.589G>T p.(Gly197Cys).
d. Additional annotative punctuation should also be spaced from the variant description rather than "touching" the description. For example, where we mark particular variants in a table to reference them in the figure legend
   - NM_000088.4:c.589G>T *
   - NM_000088.4:c.589G>T †
e. There should be no additional punctuation in the variant description. For example, the following are incorrect. Refer to the VariantValidator validation files for accurate nomenclature.
   - NC_000017.11:g.50,198,002C>A (numbers should not include commas)
   - NC_000017.11:g.50198002C > A (spaces should not be added)
f. These rules should also be applied in tables and figures.

9. Additional final formatting notes to Journal production copy editors/type setters:
   a. HGVS variant descriptions must only contain valid ASCII characters. The use of styling characters breaks the description and makes it invalid. Do not add style characters for the sake of visualization since accurate representation is essential and facilitates findability. For example, the following are invalid:
      - NM_000088.4:c.589G>T (contains a zero-width space)
      - NM_000088.4:c.589G›T (contains a zero-width space and invalid ">")
      - NM_000088.4:c.589–1G>A (modified "-" character)
      - NM_000088.4:c.642＋1del (spaces around "+" character)
      - NM_000088.4:c.589__590del (spaces around "_" character)
   b. Note: VariantValidator will identify these characters and warn of their existence and location. Paste directly into: https://variantvalidator.org/service/validate/
   c. These rules should also be applied in tables and figures.
   d. At this stage, we reiterate that tables must remain as text based and searchable by text search algorithms. Tables must not be converted into an image based format.

Technical Editor Guidance

Background

This document presents the technical editing (TE) process developed at *Genetics In Medicine* (GIM) to address widespread issues in the reporting of genomic data in biomedical journals. Technical Editors (TEs) play a vital role in ensuring that genomic variant descriptions within submitted manuscripts meet recognized nomenclature standards, particularly those set by the Human Genome Variation Society (HGVS), the International System for Human Cytogenomic Nomenclature (ISCN), and the American College of Medical Genetics and Genomics (ACMG). Their work is distinct from the traditional peer review: while peer reviewers assess scientific merit, TEs focus on the accuracy, consistency, and clarity of variant reporting.

Standardized variant nomenclature is critical for clinical diagnostics, literature searchability, and gene-disease discovery. However, GIM's internal evaluation revealed pervasive non-compliance with nomenclature standards across manuscripts, even when detailed author instructions were provided. This failure has real-world consequences—misreported variants can impede reproducibility, mislead readers, and ultimately compromise patient care.

The TE process at GIM was designed to systematically identify and correct these issues prior to publication. By implementing structured review protocols the process supports both authors and editors in improving the quality and consistency of variant reporting.

Introduction

The purpose of technical editing is to reduce the likelihood of variant nomenclature errors within published manuscripts in an effort to improve variant findability in clinical and research settings. All manuscripts being considered for publication which contain genetic variants should be subject to review by a member of the TE team. These manuscripts are initially identified by a member of the Journal's editorial staff and then distributed for technical review to a member of the TE team during the peer review process. As stated above, a TE's focus is on accurate, consistent, and clear variant reporting that adheres to the current variant nomenclature and classification standards available at the time that the manuscript was written.

As an introduction to the technical review process, technical reviewers should initially familiarise themselves with the journal's guidance to authors https://www.elsevier.com/journals/genetics-in-medicine/1098-3600/guide-for-authors (section "Terminology") and the latest publication by the HGVS variant nomenclature committee (10.1002/humu.22981). The HGVS nomenclature standard is actively updated and maintained, and the latest guidance can be found on the HGVS variant nomenclature committee website (http://varnomen.hgvs.org/). Technical reviewers should also be familiar with the ISCN standards which were updated in 2024. The latest ISCN-related guidance can be found through a subscription service: https://iscn.karger.com/. Also, be familiar with VariantValidator and how to use it (https://variantvalidator.org/,

https://doi.org/10.1002/humu.23348. Guidance can be provided by contacting the team https://variantvalidator.org/help/contact/).

Guidance on gene names, symbols, and IDs

Gene symbols and gene names can vary over time. For instance, the gene symbol *LEPRE1* was altered to *P3H1* in the mid-2010s, as the former was considered politically incorrect because of the "leprechaun" moniker found in the gene name (leucine proline-enriched Proteoglycan (leprecan) 1). This has led to discrepancies in the scientific literature, with older references using *LEPRE1* and newer ones employing *P3H1* for the same gene. To avoid such issues in the future, experts like the Vertebrate Gene Nomenclature Committee (VGNC) suggest adopting stable gene identifiers, such as "HGNC:19316." Utilizing stable identifiers alongside gene symbols and names ensures consistency and accuracy in gene identification and references in research.

Notes:

- Gene symbols should not be referred to as gene names, e.g., *P3H1* is a gene symbol and "prolyl 3-hydroxylase 1" is the gene name.
- Gene symbols must be accompanied by the HGNC gene ID *e.g.,* HGNC:19316. Although other stable identifiers are also available (e.g., the National Center for Biotechnology Information (NCBI) in the United States has assigned "64175" as the *P3H1* gene's stable identifier, and the Ensembl project has assigned "ENSG00000117385) HGNC gene IDs are required by GIM.
- OMIM IDs should accompany any diseases listed within the manuscript but are not presently accepted by GIM as stable gene identifiers.

Commented [10]: There is another recommended source too. One Heidi is in favour of. Bob, can you remember what it was?

Associating genes with transcript reference sequence IDs

In general, authors have a tendency towards stating a gene symbol *e.g., COL1A1* then describing variation in the body text of an article in the format c.1A>T. This is acceptable so long as a reference sequence is provided in the body text of the article. The association between a gene symbol and the reference sequence should be in the format *e.g., COL1A1* (NM_000088.4), and should be provided either prior to the first variant description, or in the methods section.

For articles containing multiple genes, the same convention should be followed, but additionally authors should categorize each description by also providing the gene symbol in parentheses *e.g.,* (*COL1A1*) c.1A>T (Note: this is only necessary for articles reporting variation in multiple genes). This format makes it clear which gene variants "belong to", and including an association of gene symbols with reference sequences in the methods section ensures readers can identify the correct reference sequence for each variant. When reviewing, take the time to ensure you can figure out the correct reference sequence for each variant. If you are unsure, then it is likely the readers will be unsure.

In some instances, authors may be required to describe variation in the context of more than one reference sequence for a gene. For example, *COL5A1* has 2 transcript isoforms

which utilise 2 mutually exclusive exons. In such instances, the gene and reference sequence associations should still be made; *e.g., COL5A1* (NM_000093.5, NM_001278074.1). In such instances, it is always necessary to state the exact reference sequence; *e.g.,* NM_000093.5:c.5136+68_5136+73delinsT or NM_001278074.1:c.5072_5077del and not (*COL5A1*) c.5136+68_5136+73delinsT or (*COL5A1*) c.5072_5077del.

Notes:

- Editors should check whether provided reference sequences (prefixed NM_ from RefSeq or ENST from Ensembl) are MANE Select or Mane Plus Clinical 10.1038/s41586-022-04558-8. If not, authors should be asked to update the descriptions if possible and justify not updating if it is not possible.
- LRGs (https://www.lrg-sequence.org/) and RefSeqGene IDs (NG_) are no longer supported by the HGVS Variant Nomenclature Committee but are still valid. However, authors should be encouraged to use MANE Select or MANE Plus Clinical transcripts where possible, particularly if data presented is not 3rd party.
- It is not uncommon to provide a transcript reference sequence to describe protein variation. Since the protein sequence ID is provided in the metadata of transcript reference sequence record, this is allowed.
- A gene symbol should never be included in a variant string; *e.g.,*NM_000088.3(COL1A1):c.589G>T. If authors wish to associate a gene symbol with the variant, then something like (*COL1A1*) NM_000088.3:c.589G>T is valid.
- A variant, *e.g.,* NM_007375.3:c.1144G>C p.(Ala382Pro), can be provided with a unique shorthand such as "A382P". Authors must make the association very clear at the first instance that the variant is reported and correct HGVS/ISCN nomenclature must be used before this is valid.
- Check that variants in tables are searchable using a tool such as MS Word's integrated search tool.

Guidance on reference sequences

The HUGO HGVS nomenclature committee provides extensive guidance on reference sequence types, and which should be used (http://varnomen.hgvs.org/bg-material/refseq/).

Notes:

- The LRG project is no longer maintained. Where possible, authors should be encouraged to present data in the context of MANE Select and MANE Plus Clinical transcripts.

Which variant descriptions and peripheral data must be presented?

Variants identified in Next Generation Sequencing (NGS) based studies (including Genome Sequencing (GS) Exome Sequencing (ES); gene panel sequencing; and RNA-Seq) require the alignment of sequence reads to a genome build. Authors must provide the genome build used for the analysis in the methods section. This may be GRCh37 (hg19) or GRCh38 (hg38).

We encourage the use of GRCh38 where possible. The authors should also provide the source of the genome build, *i.e.,* where they downloaded it from, with the version number. RNA-Seq data should also identify the transcriptome used and the source. Ideally this is a genbank or RefSeq genome build ID but may also be a web link to the specific genome build.

All variants must be described at the genomic level because they were identified via alignment to a genome build (this includes RNA-Seq data). These have the format "NC_000007.14:g.94409821_94409823del". We do not expect to see these descriptions in the body text, but authors should provide them in a variant table either in the manuscript or the supplement (Refer to section "Using rs numbers to describe variation").

Notes:

- Chromosome numbers are not valid reference sequence IDs, even when a genome build is provided. A genomic reference sequence (*i.e.,* NC_#) must be provided.

All variants should also be described at the cDNA (c.) level, *e.g.,* NM_000089.4:c.1035_1035+2del, and protein (p.) level *e.g.,* NP_000080.2:p.? (Note: the NP# is redundant with the NM# and is not required unless desired by authors; in addition, please refer to the section "When to use parentheses" for more p. nomenclature guidance). These descriptions will be the most prevalent in the body text of manuscripts and the formatting must be correct (Refer to sections "Associating genes with transcript reference sequence IDs", "Absent reference sequence and reference sequence selection", "Whitespace and incorrect characters within descriptions and in-between descriptions").

Variants that affect splicing, such as the one we are using in this section, may require additional descriptions to show the possible outcomes of splicing. An r. description is particularly useful for this, *e.g.,* NM_000089.3:r.1034_1036del, which also allows us to provide a p. outcome NP_000080.2:p.Val345del (Refer to the section "When to use parentheses").

Another example is NC_000010.11:g.88939507C>T. Functional studies, including RNA sequencing, suggest that this variant affects splicing but the effect is not complete, so we get 2 outcomes. The correct c. description is NM_001613.4:c.808G>A. The 2 outcomes are r.808g>a p.Gly270Arg and r.617_808del p.Ala206_Ile269del.

Notes:

- We encourage the use of the 3-letter amino acid code for p. descriptions rather than the single-letter code because it aids findability.
- Frameshifts must be described in full, *e.g.,* p.(Gly197ValfsTer68), rather than the shorthand "p.(Gly197Valfs)". However, the shortened format can be used as an alias as long as the full description is initially provided in the manuscript.
- The three-letter amino acid code for termination should be used rather than * e.g. p.(Gly197ValfsTer68) and not p.(Gly197Valfs*68)
- RNA (r.) descriptions are independent of cDNA descriptions (c.).
- RNA descriptions must only be used if the RNA has been sequenced and must not be inferred from a cDNA description.

- c. and g. descriptions provided by authors must also only be used if the DNA sequence has been confirmed.
- r. descriptions use the RNA alphabet which is lower-case and uses the base “u” instead of “t”.
- Unlike c. descriptions, r. descriptions refer to processed mRNA rather than precursor mRNA, so RNA descriptions cannot be intronic. This is why we can describe the variant NC_000007.14:g.94409821_94409823del, NM_000089.4:c.1035_1035+2del as r.1034_1036del and p.Val345del.

Note: Manuscripts which report pharmacogenetic star alleles are the exception to the requirements above. It is acceptable that such manuscripts only include the star allele nomenclature (and not the individual variants which comprise the star allele) provided a PharmVar ID (*i.e.,* PVID; see https://www.pharmvar.org/) is included alongside all star alleles at their first mention in the primary text of the manuscript as well as in all supporting tables. Manuscripts describing novel star alleles and/or those that comment on the function of specific variants within a star allele should follow the HGVS nomenclature-related guidance above.

Guidance on Structural Variant (SV) descriptions

The HGVS does not currently provide a strong syntax for describing structural variants (which includes copy number variants, *i.e.,* “CNVs”).. The del syntax does work well when used in the context of a genomic reference sequence (NC_). However, the dup syntax in the HGVS nomenclature is reserved for tandem duplication. The HUGO HGVS variant nomenclature committee are creating CNV guidelines, however, until this is developed, we will use a chromosome-based nomenclature system (ISCN 2024).

For each SV, the technique (‘arr’ or ‘seq’, for array vs sequencing, respectively), the genome build, cytobands, start and end coordinates, and number of copies must be provided. In addition, the CNV size must be provided (in bp, kb, or Mb; ideally within a variant table), and clinically relevant genes within the deleted or duplicated interval should be listed. Please note that stable gene identifiers (*i.e.,*, HGNC) should also be provided for these genes.

Some common examples of technology-specific SV nomenclature in accordance with ISCN 2024 guidelines can be found below. However, TEs should note that this is not comprehensive and should reference the ISCN 2024 for additional guidance.

CNVs detected by microarray should be in the format:

arr[GRCh38] 1p36.33p36.32(827048_3736354)x3 which indicates a gain of 1p36.33p36.32.

arr[GRCh38] 1p36.33p36.32(827048_3736354)x1 which indicates a loss of the same region.

CNVs detected by sequencing should have both the ISCN-like description and the HGVS-like nucleotide variant descriptions:

For example, an interstitial deletion should be represented in the following two ways (please note that the ‘and’ in the examples below below should not be included but was added to highlight the difference between the ISCN-like and HGVS-like nomenclature):

seq[GRCh38] del(X)(q21.31q22.1) and NC_000023.11:g.89555676_100352080del

Commented [11]: Peter, do we require CNV variant classification based on the 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) https://pmc.ncbi.nlm.nih.gov/articles/PMC7313390/?

Commented [12]: We should yes. Bob, can we put this in?

Commented [13]: Great catch Mari! I added in the reference/this requirement alongside the Richards et al SNV scoring criteria paper when it is mentioned above.

Peter, we do not presently have a subsection for variant interpretations outside of the checklist above. Are you ok with that or would you like to see a section about variant classifications in the main part of these TE guidelines?

Commented [14]: I think we should have a section in the main guidelines as well yes.

Whereas a duplication should be represented as:

seq[GRCh38] dup(X)(q21.31q22.1) and NC_000023.11:g.89555676_100352080dup

Please note that both the ISCN-like and HGVS-like nomenclature are especially necessary for duplications, as the ISCN-like nomenclature is non-specific (i.e., does not indicate whether the duplication is happening in-tandem or inserted elsewhere in the genome).

For example, if the duplication above was not in-tandem but was inserted into a different location, it would be written as follows using 'ins' rather than 'dup' within the HGVS-like nomenclature:

seq[GRCh38] dup(8)(q24.21q24.21) and NC_000008.11:g.128749160_128749161ins128746677_128749160 (describes an ~2.4kb duplicated sequence within chromosome 8 that is inserted between nucleotides 128,749,160 and 128,749,161).

If the duplicated sequence is inserted and inverted, then 'ins' and 'inv' should be included:

seq[GRCh38] dup(8)(q24.21q24.21) and NC_000008.11:g.128746676_12874677ins128746677_128749160inv describes an ~2.4kb duplicated sequence within chromosome 8 that is inserted in an inverted orientation between nucleotides 128,746,676 and 12,874,677.

Inheritance of CNVs

If the inheritance of the CNV is known, 'mat' or 'pat' (for 'maternal' and 'paternal', respectively) should be provided at the end of the ISCN-like description. For example: seq[GRCh38] dup(X)(q21.31q22.1) mat indicates maternal inheritance of this interstitial duplication of chromosome X.

Breakpoints of CNVs

If the exact breakpoints of the CNV are not known (e.g., CNV calling from exome data or targeted sequencing), this uncertainty should be reflected in the CNV nomenclature by providing the exact coordinates of the data captured:

seq[GRCh37] 12q21.32q21.32(88928628x2,88939469_88939743x1,88973993x2) and NC_000012.11:g.(88928629_88939469)_(88939743_88973992)del

Aneuploidy Detected by Sequencing

Aneuploidy that is detected by sequencing without structural resolution by chromosome analysis can be represented as ISCN or HGVS nomenclature:

ISCN: seq(13) x3
HGVS: NC_000013.11:g.(pter)_(qter)[2] - indicating that there is one additional copy of chromosome 13 sequences detected

Aneuploidy that is detected by sequencing and chromosome analysis can be described using either ISCN or HGVS nomenclature:

ISCN: 47,XY,+13.seq (13)x3

HGVS: NC_000013.11:g.(pter)_(qter)sup

Common nomenclature errors

Absent reference sequence and reference sequence selection

HGVS nomenclature clearly states that all variation must be defined in the context of a reference sequence. Detailed guidance about reference sequences that are considered suitable for reporting can be found on the variant nomenclature committee website. http://varnomen.hgvs.org/bg-material/refseq/. Authors must provide an appropriate reference sequence for all variant descriptions. Recommended practice is to provide the reference sequence in the methods section of articles, but it is also useful to include the reference sequence the first time variation in the context of the chosen reference sequence is described (refer to "Associating genes with transcript reference sequence IDs").

Whitespace and incorrect characters within descriptions and in-between descriptions

A "full and complete" HGVS description has the format NM_000088.4:c.1A>T. There should be no whitespace in the description, and not the colon before the reference sequence type identifier, in this case :c. The format NM_000088.4(COL1A1):c.1A>T is common, but is not standards compliant. If authors wish to use the gene symbol, they should specify the reference sequence and gene symbol, as above *e.g.,* NM_000088.4 (*COL1A1*), then use the format (*COL1A1*) c.1A>T (note the space and the lack of a colon). There should never be any other whitespace in the descriptions *e.g.,* c.1A > T is not valid.

Commented [15]: Are we also OK with "COL1A1 (NM_000088.4)" in associating the gene ID with the transcript number before any variants are described? I see this in many papers.

Authors have a tendency to try and "merge" descriptions together using a delimiter e.g., NM_000088.4:c.1A>T/p.(Met1?) or NM_000088.4:c.1A>T;p.(Met1?). This is not correct and should be replaced with white space i.e., NM_000088.4:c.1A>T p.(Met1?). The rule of thumb is that each description is a separate entity and should be treated as such.

When to use parentheses

Note: This section is under review and may be deprecated in the next version of these guidelines. However, we currently recommend and enforce strict adherence to the HGVS nomenclature standard.

The HGVS nomenclature uses parentheses to denote a "prediction". For example, if a variant is described as NC_000017.11:g.50201513T>A NM_000088.4:c.1A>T, the p. description is inferred from the c. so is a prediction, *i.e.,* NP_000079.2:p.(Met1?). If, however, the RNA has been sequenced (this only applies for sequencing of the RNA which is not aligned to a genomic reference sequence *e.g.,* RNA-Seq) we can describe a variant as NM_000088.3:c.589g>u (Note: The RNA alphabet is lower-case and u replaces t) and NP_000079.2:p.Gly197Cys.

This convention can be very useful for describing the consequence of variants that affect splicing. For example, NC_000007.14:g.94409821_94409823del is described at the c. level

as NM_000089.3:c.1035_1035+2del (*COL1A2*). The variant is described at its most 3 prime position and in this case, the variation at the c. level spans the last base of an exon and the first 2 bases of the subsequent intron. When described as an RNA variant, the transcript is considered to be “mature mRNA” so introns are not considered, and the variant is correctly described as NM_000089.3:r.1034_1036del and NP_000080.2:p.Val345del (no parentheses because the RNA sequence is confirmed).

Commented [16]: Under review pending feedback from the ACMG committee

Using external annotations to describe variation

A "single-nucleotide variant" (SNV) refers to a single-nucleotide change, which is the most abundant form of genomic variation and widely prevalent. Instead of "SNV," the term "single-nucleotide polymorphism" (SNP, pronounced as "snip") is often used, but it is no longer recommended in the context of human genetics because it might imply a benign effect. Similarly, the term "mutation" is no longer preferred for germline sequence variation as it might suggest a pathogenic effect. To avoid misunderstandings, the term "SNV" is preferred, and SNVs are classified as "pathogenic,", “likely pathogenic,” “of uncertain significance,” “likely benign,” or"benign." To facilitate the identification and classification of SNVs, the National Center for Biotechnology Information (NCBI) established the public-domain database called dbSNP. This database contains records on various variant types, including deletions, duplications, inversions, and deletion insertions with fewer than 50 base pairs. Each entry in dbSNP is assigned an "rs number," which the NCBI defines as a "RefSNP number." However, using rs numbers alone is not sufficient to describe SNVs because they refer to genomic coordinates rather than individual variants. Therefore, multiple variants may be identified by the same rs number in dbSNP records.

To accurately describe individual genetic variants, dbSNP records contain terms following the Human Genome Variation Society's format. Manuscripts using rs numbers must also provide the relevant genomic HGVS description to avoid ambiguity. This comprehensive approach ensures that the information regarding SNVs in dbSNP remains accurate and specific to individual variants.

A key benefit of external annotations is that they often contain additional relevant data. dbSNP IDs are useful since they link to several external resources provided by NIH and EBI. However, they are not necessarily the most useful annotations for use clinically. Recommend to authors that they use one of (where relevant)

- ClinGen allele registry ID (highly recommended)
- dbSNP ID
- ClinVar ID or LOVD ID (highly recommended)

Termination codes

There are 2 ways to represent a termination in a p. variant: the single letter amino acid code “*” *e.g.*, NP_065202.2:p.(V87*) or the three-letter “Ter” *e.g.,* NP_065202.2:p.(Val87Ter). The format 'NP_065202.2:p.(Val87*) has historically been acceptable, but we have found

that it affects the findability of variants. Therefore, we do not support this syntax and insist that authors use Ter when describing amino acid terminations in the 3-letter amino acid alphabet. "*" must only be used when the single-letter amino-acid alphabet is being utilized throughout the manuscript. Additionally, p.(Val87X) is not correct because "X" is the IUPAC symbol for any amino acid.

Uncertain positions and the allele syntax

The uncertain position syntax is somewhat complex and ambiguous. They are also hard to validate, such is the case with uncertainty. Familiarize yourself with the uncertain description formats http://varnomen.hgvs.org/recommendations/uncertain/. If you require assistance, contact the VariantValidator team.

The allele syntax is less complex. Guidance can be found here http://varnomen.hgvs.org/recommendations/DNA/variant/alleles/.

Notes:

- Allele syntax should only be used to describe 2 or more variants on the same reference sequence.
- Allele syntax should not be used to describe the outcomes of a single variant using the r. syntax.
- Allele syntax should not be used to describe a single variant c. and p. level.
- The use of brackets (*i.e.,* "[ ]") is discouraged in any variant nomenclature strings unless they are being used to accurately convey variant information using the HGVS standard, for example to indicate alleles.

Technical Editor Description

Main Tasks:
For articles submitted for publication in *GIM* and *GIM Open*:

- To ensure the DNA sequence variants conform to current HGVS nomenclature and ISCN nomenclature.
- To set CNV nomenclature standards [with David Miller and Bo Yuan as consultants] and ensure CNV variants conform to those journal standards
- To ensure small sequence variants are classified according to the ACMG/AMP standards[1] and that CNVs are classified according to the ACMG/ClinGen[2] standards, or authors have detailed the classification system they have used and why these classifications are not possible. With gene discovery manuscripts, the classification should be assigned as if there is a known gene-disease association with this process described briefly in methods and the inclusion of a gene-disease scoring matrix[3]. In addition to the classification, the variant scoring results (both individual scoring criteria and overall variant score/assessment) should be listed in a table (this can be supplementary and should cross reference the variant table (see and link to "Presentation of variants in GIM")).

- Resolve any queries around SNV, CNV, or other variant type nomenclature or variant classification during the peer review process.

Job Requirements

- PhD and/or MD with a genetics/genomics focus.
- Understanding of DNA sequence and copy number variant nomenclature and classifications.
- LGG certification is desirable.

GIM | GIMO Instructions for Authors (Online)

Introduction

These authorship guidelines are designed to support the accurate and consistent reporting of genomic data in manuscripts submitted to Genetics in Medicine (GIM) and GIM Open. Clear, standardized variant descriptions and classifications enhance the findability, interoperability, and clinical utility of published work. By following internationally accepted standards, authors contribute to the reliability of genomic data in the scientific literature, enabling reproducibility, effective data sharing, and broader impact in research and patient care.

Description of Genetic Disorders

Genetic disorders should be accompanied by their corresponding OMIM numbers if available. For details on OMIM numbers, please refer here. Otherwise, please include other suitable disease identifiers e.g., MONDO disease term.

Commented [17]: I think you may need to add this above, I had asked a question about it, because I didn't think the guidelines were clear.

Commented [18]: I added a line about this above today. Please let me know if you think that it should be further emphasized (perhaps by carving out a specific subsection rather than a brief mention?)

Genome build nomenclature

Studies which identify novel variants (*e.g.*, Genome Sequencing or RNA Sequencing studies) must include a genome build in the methods section of the manuscript. We recommend that studies use the GRCh38 genome build, however, we are aware that use of GRCh37 is still common.

The genome build should be referred to as GRCh38 (or hg38) or GRCh37 (of hg19). Additionally, an ID and version number for the genome build should be provided such as a GenBank ID (e.g. GCA_000001405.29) or RefSeq ID (e.g. GCF_000001405.40) or a unique identifier or web link for genome builds obtained from alternative sources e.g. the GATK GRCh38 resource bundle (https://console.cloud.google.com/storage/browser/genomics-public-data/resources/broad/hg38/v0;tab=objects?inv=1&invt=Ab0nQQ&prefix=&forceOnObjectsSortingFiltering=false). The resource from which the genome build was obtained should be referenced in the methods section of the manuscript.

Gene Nomenclature Requirements

Approved gene symbols and nomenclature can be obtained from the HUGO Gene Nomenclature Committee (HGNC). Authors must include the approved gene symbol and corresponding HGNC identifier in the manuscript text, for example *P3H1* (HGNC:19316). The gene symbol must not be confused with the gene name e.g., "*P3H1*" is the gene symbol and

"prolyl 3-hydroxylase 1" is the gene name. The ID provides a stable reference for the gene since gene symbols and gene names can change over time. Human gene symbols and loci are given in italicized capital letters and Arabic numbers. Protein products are not italicized.

Mouse gene/variant nomenclature should use the International Committee on Standardized Genetic Nomenclature for Mice.

Nomenclature of DNA Sequence Variants and Variant Data Sharing

Documenting variation in our genomes is an important undertaking for human research and clinical care. Accuracy in the notation of DNA sequence variants is essential for the success of this endeavor. Accurately naming and deposition of variants in appropriate databases ensures that the clinical literature is discoverable and usable by other researchers.

Because of the importance of the issue and the consensus on the rules *Genetics in Medicine* (*GIM*) require compliance with the Human Genome Organization (HUGO) Human Genome Variation Society (HGVS) variant nomenclature committee recommendations for describing sequence variants and compliance with the International System for Human Cytogenomic Nomenclature (ISCN) before manuscripts can be accepted and published.

GIM also strongly encourages all authors to submit DNA sequence variants reported in manuscripts to a public database (e.g., ClinVar, Global Variome, or shared LOVD) prior to publication and to provide the resulting database entry ID in the manuscript for each variant.

MANDATORY: Compliance with HGVS and ISCN nomenclature

DNA sequence variant nomenclature should follow current recommendations of the HGVS Variant Nomenclature Committee and CNV/structural variant nomenclature should follow current recommendations of the ISCN committee. ***Acceptance and/or publication of your manuscript will be delayed if authors do not follow these guidelines***. Please visit HGVS and https://iscn.karger.com/ for the latest nomenclature updates, for examples of acceptable nomenclature, guidance concerning reference sequences, or if you have further questions. The following section provides instructions for the presentation of variants in *GIM*.

- The use of legacy nomenclature is not permitted without being associated with an accurate HGVS description of the legacy term.
- When variant descriptions are reproduced from an existing source e.g. previous publication or database entry, it is the author's responsibility to re-validate the

Commented [21]: Would prefer that we incorporate this into a more general nomenclature section below. We expect that authors publishing SNVs will appropriately follow HGVS guidance without listing examples and I think it's ok to expect that authors publishing CNVs will appropriately follow ISCN guidance without listing examples.

Commented [22]: Volunteers please :)

Deleted: Human Chromosome Nomenclature including CNV nomenclature and Structural Variation nomenclature

Deleted: ¶

Deleted: *GIM* follows the International System for Human Cytogenomic Nomenclature (ISCN) for cytogenomic variations. Many users will have institutional access through this link: https://iscn.karger.com/.

Deleted: ¶

Deleted: ¶

Deleted: The HGVS does not currently provide a strong syntax for describing large CNVs. The del syntax does work well when used in the context of a genomic reference sequence (NC_). However, the dup syntax in the HGVS ... [1]

Deleted: For each CNV, th...e technique ('arr' or 'seq', ... [2]

Deleted: ¶

Deleted: CNVs detected by Microarray should be in the ... [3]

Deleted: ¶

Deleted: arr[GRCh38] 1p36.33p36.32(827048_373635 ... [4]

Deleted: arr[GRCh38] 1p36.33p36.32(827048_373635 ... [5]

Deleted: ¶

Deleted: CNVs detected by Sequencing should have bo ... [6]

Deleted: ¶

Deleted: For example, an interstitial deletion should b ... [7]

Deleted: seq[GRCh38] del(X)(q21.31q22.1) and ... [8]

Deleted: Whereas a duplication should be represented ... [9]

Deleted: seq[GRCh38] dup(X)(q21.31q22.1) and ... [10]

Deleted: Please note that both the ISCN-like and HGV ... [11]

Deleted: For example, if the duplication above was n ... [12]

Deleted: seq[GRCh38] dup(8)(q24.21q24.21) and ... [13]

Deleted: If the duplicated sequence is inserted and in ... [14]

Deleted: seq[GRCh38] dup(8)(q24.21q24.21) and ... [15]

Deleted: ¶

Deleted: Inheritance of CNVs...¶ ... [16]

Deleted: If the inheritance of the CNV is known, 'mat ... [17]

Deleted: Breakpoints of CNVs...¶ ... [18]

Deleted: If the exact breakpoints of the CNV are not k ... [19]

Deleted: seq[GRCh37] ... [20]

Deleted: ... [21]

nomenclature, updating it as required, and to provide a citation / database ID for the source of the previous report.

- All CNVs must be accompanied by a list of disease-associated genes (with MIM numbers) that are overlapped within the CNV interval as well as the size of the CNV.

Note: Manuscripts which report pharmacogenetic star alleles are the exception to the requirements above. It is acceptable that such manuscripts only include the star allele nomenclature (and not the individual variants which comprise the star allele) provided a PharmVar ID (*i.e.,* PVID; see https://www.pharmvar.org/) is included alongside all star alleles at their first mention in the primary text of the manuscript as well as in all supporting tables. Manuscripts describing novel star alleles and/or those that comment on the function of specific variants within a star allele should follow the HGVS nomenclature-related guidance above.

MANDATORY: Variant pathogenicity assessment

All variants reported within submitted manuscripts should be assessed for pathogenicity in accordance with current SNV and CNV interpretation guidelines, including gene-specific guidelines, when applicable. Please note that this information may appear in a supplemental document and should include the scoring criteria (e.g., "PVS1" or "2A = 1") and final classification (e.g., "Likely Pathogenic") for each variant reported.

Authors reporting variants within genes that are not yet established as having a gene-disease relationship should still classify variants as if the gene-disease relationship is already established and note that this approach was taken in the text.

In addition, manuscripts that are working to establish a novel gene-disease relationship should evaluate the strength of the established gene-disease association framework and include supporting documentation within their manuscript.

Amino Acid Three Letter Codes

Our preference is that you follow International Union of Biochemistry and Molecular Biology guidelines on the three letter codes that should be used to describe amino acids.

- Termination codes must match the amino-acid alphabet used in the document i.e. Ter must be used when utilising the 3-letter amino acid code, and * when using the single letter amino-acid code
- Frameshifts must be described in full e.g. p.(Gly197ValfsTer68) rather than p.(Gly197Valfs) however, the latter can be used as a shorthand if the full description is also provided
- Parentheses must be used if the variant protein sequence or the variant RNA sequence has not confirmed by direct sequencing

Commented [23]: Under review

HGVS Nomenclature Validation

All instances of variants, in the manuscript and tables, **must** be verified using variant description validation software, VariantValidator.

VariantValidator provides two ways to validate variant descriptions depending on the number of variants you have to validate:

Submission of up to 10 variants
Input each variant as prompted on the webpage then find the "Print PDF" button under the "Recommended Variant Descriptions" table. Save a copy of each PDF and upload each PDF with your submission.

Submission of 10 or more variants
Use the batch tool. Instructions and help for this tool can be found here. Please upload the results with your submission as an Excel table that includes all the variants.

Errors and Warnings

If you receive errors or warnings, these are provided as guidance to authors and editors to help you report your variants according to the recommendations. Please refer to https://variantvalidator.org/help/instructions/ for a list of potential warnings from VariantValidator.

Please contact the software systems admin to resolve the errors or warnings prior to submitting the variant file to the journal.

Contact Information

VariantValidator: https://variantvalidator.org/help/contact/. Or join the discussion on https://app.gitter.im/#/room/#variantValidatorDiscuss_community:gitter.im where you can report issues, discuss variants or even request features.

This tool is freely available on the web and relies on funding from the host universities, grant income and community contribution. Please support these tools by citing them in your manuscript if they have been helpful or by making a contribution to their chosen charities and providing a supportive comment. These contributions really help when it comes to securing funding and helps keep the software free.

Citing VariantValidator

1. Freeman, PJ, Hart, RK, Gretton, LJ, Brookes, AJ, Dalgleish, R. VariantValidator: Accurate validation, mapping and formatting of sequence variation descriptions. *Human Mutation*. 2018;39:61-68. https://doi.org/10.1002/humu.23348.

2. Freeman, P.J., Wagstaff, J.F., Fokkema, I.F.A.C. et al. Standardizing variant naming in literature with VariantValidator to increase diagnostic rates. Nat Genet 56, 2284–2286 (2024). https://doi.org/10.1038/s41588-024-01938-w

Presentation of Variants in GIM

When reporting a variation, it is important that the identifying variant be described; all other variant descriptions are considered annotative. Authors **must** include a relevant variant table in the supplement text of each manuscript, selecting the most appropriate from the below criteria:

- For variants identified by NGS sequencing techniques, the genomic description (g.) must be provided along with the HGVS c. and, where relevant, p. descriptions. This includes RNA Seq methodologies where sequences are aligned to a reference genome rather than a reference transcript.
    a. Gene descriptions, e.g., RefSeqGene or LRG are no longer recommended by the HGVS gene nomenclature committee. Instead, authors should refer to HGVS guidance on the use of MANE Select and MANE Plus Clinical transcripts. Although MANE Select and MANE Plus Clinical transcripts are preferred, justification for the use of an alternate transcript may be provided within the methods section of the manuscript.
    b. This prevents spurious reporting in the context of a reference sequence which was not used to identify variation, but was inferred from mapping via a transcript reference sequence.
- For variants identified by RNA/cDNA sequencing, an RNA level (r.) description may be provided
    a. In such cases the g. and c. description should not be provided unless they have been confirmed at the DNA level e.g., the genomic variant is confirmed by DNA sequencing methodologies
- For structural variants (including CNVs), nomenclature guidance within ISCN 2024 related to whether ISCN, HGVS, or a both variant descriptions are warranted should be followed.

For HGVS-related examples we direct the authors to: https://varnomen.hgvs.org/

For ISCN-related nomenclature updates in the 2024 version of the guidelines, we direct the authors to: https://karger.com/cgr/article/165/1/1/924566/ISCN-2024-Summary-of-Revisions-and-New

Other important considerations

Reference sequences defined in the HGVS nomenclature guidelines must be used for

reporting sequence variants. Authors should always include the Accession Number of the relevant reference sequence(s), with version number where applicable (e.g.: RefSeq NM_003002.3, or GenBank NC_000011.10), in the Materials and Methods section and as a footnote in any tables listing variants. Please note: RefSeq or/and Ensembl transcript reference sequences may be used. We recommend the use of transcript references that have been denoted as the default reporting references through the Matched Annotation from the NCBI and EBI (MANE) project, or alternative transcripts denoted as Mane Plus Clinical. may be used. once approved by the HGVS variant nomenclature working group.

If alternative nomenclature schemes are commonly found in the literature, they may also be used in addition to approved nomenclature, but they must be clearly defined.

Standard HGVS nomenclature using g. annotation and identifying the genome build must be used for non-coding variation, including those variants identified in GWAS studies (e.g., NC_000017.11:g.50201450C>T). Variants may also be described using dbSNP genomic location identifiers, *in addition to approved nomenclature*, if the specific nucleotide change is also included. The use of gnomAD style descriptions e.g. chr17-.50201450-C-T is permitted for use as a shorthand, so long as the HGVS g. description is also provided. This includes use in tables where both descriptions must appear.

Author Checklist

- *Standardised disease names have been provided including an identifier (e.g. MIM ID)*
- *HGNC approved gene symbols are provided and italicised*
- *HGNC gene IDs are provided for each gene symbol*
- *Where relevant, a genome build for the analysis including a genbank or RefSeq identifier for the build (preferred) or a web link to the genome build. The correct format for describing the build is e.g. GRCh38 or hg38.*
- *ISCN descriptions for copy number and structural variation*
- *HGVS descriptions of each variant at the level of g. c. and p. (note, the p. is not required for intronic or intergenic variants)*
- *Reference sequences must be provided for each description*
- *Termination codes must match the amino-acid alphabet used in the document i.e. Ter must be used when utilising the 3-letter amino acid code, and * when using the single letter amino-acid code*
- *Frameshifts must be described in full e.g. p.(Gly197ValfsTer68) rather than p.(Gly197Valfs) however, the latter can be used as a shorthand if the full description is also provided*

- *Parentheses must be used if the variant protein sequence or the variant RNA sequence has not confirmed by direct sequencing*
- *Every document must be self-contained i.e. every variant must be described at each of the above levels and must not rely on additional documents for any other description*
- *Up to 5 variants can be included in the body text of the document. More than 5 variants we recommend the use of a standardised variant table*
- *Key variants must be classified using an accepted standard e.g. the ACMG standard. The standard used must be stated and the classification must match the terms used in the standard. In addition, the scoring used to reach the final classification must be included e.g. PP3, PVS1…*
- *Where applicable, Sequence Ontology Terms should be used to describe information which cannot be inferred from the variant descriptions.*
- *It is also recommended that DNA sequence variants reported in manuscripts be submitted to a public database e.g. ClinVar or the Leiden Open Variation Database prior to publication. (IDs can be obtained under embargo, see the ClinVar and LOVD manuals for guidance)*

Commented [24]: We need to discuss this in the group since ACMG is dropping this requirement. Pushing the HUGO HVNC to drop the use of parentheses. I support the ACMG decision.

Commented [25]: This is being created for the ACMG committee once finalised, I will add a copy here.

Commented [26]: https://docs.google.com/spreadsheets/d/1e1tbHkaogwA5L2Rulsc8GR3bnmy-Cj1AYgPzlUXnskU/edit?usp=sharing

Sequence Ontology (SO)

A structured, controlled vocabulary designed to describe the features of biological sequences and the impact of genetic variation on those features. SO provides standardized terms that categorize how sequence changes affect specific genomic elements like coding regions, introns, exons, splice sites, and regulatory domains. We encourage authors to apply SO terms to convey the consequences of variation in a consistent and biologically meaningful way when in circumstances where essential information cannot be extrapolated from a variant description. For example, when referring to diseases associated with imprinting a HGVS or ISCN variant description should be accompanied by a SO term and ID such as paternally_imprinted (SO:0000136) or maternally_imprinted (SO:0000135) to accurately define the origin of the imprinted region. Other use cases include, but are not limited to, short_tandem_repeat_contraction [SO:0002163] short_tandem_repeat_expansion [SO:0002162], cryptic_splice_acceptor (SO:0001570) and cryptic_splice_donor [SO:0001571].

Classification of Variants

We expect articles submitted to an official journal of the ACMG will adhere to standards and guidelines promulgated in the ACMG/AMP classification of sequence variants (Richards et al, Genet Med. 2015 May;17(5):405-424) and the ACMG/ClinGen consensus statement on

reporting copy-number variants (Riggs et al, Genet Med. 2020 Feb;22(2):245-257). For research reporting classification of newly described variants or reclassification of variants, we require authors to describe the variants using the ACMG/AMP classification system. In addition to the classification, the Variant scoring results should be listed in a table (this can be supplementary, and should cross reference the variant table (see and link to "Presentation of variants in GIM ")).

For manuscripts reporting previously classified variants whose classification was assigned using a different standard, we will accept that classification with proper documentation and explanation in the methods section. However, including the ACMG/AMP classification of that sequence variant alongside the previous classification is strongly encouraged. We are aware that other classification standards exist and in special cases with prior discussion with the editorial office, newly described variants may be classified using other standards, but generally with ACMG/AMP classification also listed. Please contact the editorial office at geneticsinmedicine@acmg.net for questions or further details.

Page 18: [1] Deleted	**Lisa Lansdon**	**24/07/2025 21:15:00**
Page 18: [1] Deleted	**Lisa Lansdon**	**24/07/2025 21:15:00**
Page 18: [2] Deleted	**Lisa Lansdon**	**24/07/2025 21:15:00**
Page 18: [2] Deleted	**Lisa Lansdon**	**24/07/2025 21:15:00**
Page 18: [2] Deleted	**Lisa Lansdon**	**24/07/2025 21:15:00**
Page 18: [3] Deleted	**Lisa Lansdon**	**24/07/2025 21:15:00**
Page 18: [3] Deleted	**Lisa Lansdon**	**24/07/2025 21:15:00**
Page 18: [3] Deleted	**Lisa Lansdon**	**24/07/2025 21:15:00**
Page 18: [4] Deleted	**Lisa Lansdon**	**24/07/2025 21:15:00**
Page 18: [4] Deleted	**Lisa Lansdon**	**24/07/2025 21:15:00**
Page 18: [5] Deleted	**Lisa Lansdon**	**24/07/2025 21:15:00**
Page 18: [5] Deleted	**Lisa Lansdon**	**24/07/2025 21:15:00**
Page 18: [6] Deleted	**Lisa Lansdon**	**24/07/2025 21:15:00**
Page 18: [6] Deleted	**Lisa Lansdon**	**24/07/2025 21:15:00**
Page 18: [7] Deleted	**Lisa Lansdon**	**24/07/2025 21:15:00**
Page 18: [7] Deleted	**Lisa Lansdon**	**24/07/2025 21:15:00**
Page 18: [8] Deleted	**Lisa Lansdon**	**24/07/2025 21:15:00**
Page 18: [8] Deleted	**Lisa Lansdon**	**24/07/2025 21:15:00**

Page 18: [9] Deleted	**Lisa Lansdon**	**24/07/2025 21:15:00**
Page 18: [9] Deleted	**Lisa Lansdon**	**24/07/2025 21:15:00**
Page 18: [10] Deleted	**Lisa Lansdon**	**24/07/2025 21:15:00**
Page 18: [10] Deleted	**Lisa Lansdon**	**24/07/2025 21:15:00**
Page 18: [10] Deleted	**Lisa Lansdon**	**24/07/2025 21:15:00**
Page 18: [11] Deleted	**Lisa Lansdon**	**24/07/2025 21:15:00**
Page 18: [11] Deleted	**Lisa Lansdon**	**24/07/2025 21:15:00**
Page 18: [12] Deleted	**Lisa Lansdon**	**24/07/2025 21:15:00**
Page 18: [12] Deleted	**Lisa Lansdon**	**24/07/2025 21:15:00**
Page 18: [13] Deleted	**Lisa Lansdon**	**24/07/2025 21:15:00**
Page 18: [13] Deleted	**Lisa Lansdon**	**24/07/2025 21:15:00**
Page 18: [14] Deleted	**Lisa Lansdon**	**24/07/2025 21:15:00**
Page 18: [14] Deleted	**Lisa Lansdon**	**24/07/2025 21:15:00**
Page 18: [15] Deleted	**Lisa Lansdon**	**24/07/2025 21:15:00**
Page 18: [15] Deleted	**Lisa Lansdon**	**24/07/2025 21:15:00**
Page 18: [16] Deleted	**Lisa Lansdon**	**24/07/2025 21:15:00**
Page 18: [16] Deleted	**Lisa Lansdon**	**24/07/2025 21:15:00**
Page 18: [17] Deleted	**Lisa Lansdon**	**24/07/2025 21:15:00**

Page 18: [17] Deleted	**Lisa Lansdon**	**24/07/2025 21:15:00**
Page 18: [18] Deleted	**Lisa Lansdon**	**24/07/2025 21:15:00**
Page 18: [18] Deleted	**Lisa Lansdon**	**24/07/2025 21:15:00**
Page 18: [19] Deleted	**Lisa Lansdon**	**24/07/2025 21:15:00**
Page 18: [19] Deleted	**Lisa Lansdon**	**24/07/2025 21:15:00**
Page 18: [20] Deleted	**Lisa Lansdon**	**24/07/2025 21:15:00**
Page 18: [21] Deleted	**Lisa Lansdon**	**24/07/2025 21:15:00**