

Verifying nomenclature of DNA variants in submitted manuscripts: Guidance for journals.

Title: Verifying nomenclature of DNA variants in submitted manuscripts: Guidance for journals.

Publication year: 2021

Journal: Human mutation

Authors: American College of Medical Genetics and Genomics, Bethesda, Maryland, USA.

Society: American College of Medical Genetics and Genomics

Origin country: United States

Target countries: International

DOI: 10.1002/humu.24144

PMID: 33252176

Source: <https://pubmed.ncbi.nlm.nih.gov/33252176/>

Hum Mutat Hum Mutat 319 nihpa 9215429 Human mutation 1059-7794 1098-1004
pmc-is-collection-domain yes pmc-collection-title NIHPA Author Manuscripts PMC7961887
PMC7961887.1 7961887 7961887 NIHMS1678261 33252176 10.1002/humu.24144
NIHMS1678261 NIHPA1678261 1 Article Verifying nomenclature of DNA variants in submitted
manuscripts: Guidance for journals Higgins Jan <http://orcid.org/0000-0002-5342-382X> 1 Dagleish
Raymond 2 den Dunnen Johan T. 3 Barsh Greg 4 5 Freeman Peter J. 6 Cooper David N. 7 Cullinan
Sara 8 Davies Kay E. 9 Dorkins Huw 10 11 Gong Li 12 Imoto Issei 13 Klein Teri E. 12 Korf Bruce
14 Misra Adya 15 16 Paalman Mark H. <http://orcid.org/0000-0002-2568-0394> 17 Ratzel Sarah 8
Reichardt Juergen K. V. 18 Rehm Heidi L. 19 20 21 Tokunaga Katsushi 12 22 Weck Karen E. 23 24
Cutting Garry R. 25 1 American College of Medical Genetics and Genomics, Bethesda, Maryland,
USA 2 Department of Genetics and Genome Biology, University of Leicester, Leicester, UK 3
Human Genetics and Clinical Genetics, Leiden University Medical Center, Leiden, The Netherlands
4 HudsonAlpha Institute for Biotechnology, Huntsville, Alabama, USA 5 Department of Genetics,
Stanford University, Stanford, California, USA 6 Division of Informatics, Imaging & Data Science,
Faculty of Biology, Medicine and Health, The University of Manchester, Manchester, UK 7 Institute
of Medical Genetics, School of Medicine, Cardiff University, Cardiff, UK 8 American Society of
Human Genetics, Rockville, Maryland, USA 9 Department of Physiology, Anatomy and Genetics,
MDUK Oxford Neuromuscular Centre, University of Oxford, Oxford, UK 10 St Peter's College,
University of Oxford, Oxford, UK 11 Department of Clinical Genetics, University Hospitals of
Leicester NHS Trust, Leicester, UK 12 Department of Biomedical Data Science, Stanford
University, Stanford, California, USA 13 Division of Molecular Genetics, Aichi Cancer Center
Research Institute, Nagoya, Japan 14 Department of Genetics, University of Alabama at
Birmingham, Birmingham, Alabama, USA 15 Public Library of Science, San Francisco, California,
USA 16 Public Library of Science, Cambridge, UK 17 Cell and Molecular Biology Research, Wiley
Periodicals LLC, Hoboken, New Jersey, USA 18 Australian Institute of Tropical Health and
Medicine, James Cook University, Smithfield, Queensland, Australia 19 Medical & Population
Genetics Program and Genomics Platform, Broad Institute of MIT and Harvard, Cambridge,
Massachusetts, USA 20 Department of Pathology, Harvard Medical School, Boston,
Massachusetts, USA 21 Center for Genomic Medicine, Massachusetts General Hospital, Boston,
Massachusetts, USA 22 Genome Medical Science Project, National Center for Global Health and
Medicine, Tokyo, Japan 23 Department of Pathology and Laboratory Medicine, University of North
Carolina at Chapel Hill, Chapel Hill, North Carolina, USA 24 Department of Genetics and Medicine,
School of Medicine, University of North Carolina, Chapel Hill, North Carolina, USA 25 Department
of Genetic Medicine, McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University
School of Medicine, Baltimore, Maryland, USA Correspondence : Garry R. Cutting, Department of
Genetic Medicine, McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University
School of Medicine, Miller Research Bldg, Rm 559, 733 N. Broadway, Baltimore MD 21205, USA.
gcutting@jhmi.edu 10 12 2020 1 2021 42 1 377598 3 7 04 03 2021 16 03 2021 16 03 2021 27 04
2025 This file is available for text mining. It may also be used consistent with the principles of fair
use under the copyright law. Documenting variation in our genomes is important for research and

clinical care. Accuracy in the description of DNA variants is therefore essential. To address this issue, the Human Variome Project convened a committee to evaluate the feasibility of requiring authors to verify that all variants submitted for publication complied with a widely accepted standard for description. After a pilot study of two journals, the committee agreed that requiring authors to verify that variants complied with Human Genome Variation Society nomenclature is a reasonable step toward standardizing the worldwide inventory of human variation.

Documenting variation in our genomes is important for research and clinical care. Accuracy in the description of DNA variants is therefore essential (Tack et al., 2016 ; Yen et al., 2017). To address this issue, the Human Variome Project (HVP; <https://www.humanvariomeproject.org/>) under the auspices of Global Variome (GV) convened a committee (HVP/GV Reporting of Sequence Variants Working Group) to evaluate the feasibility of requiring authors to verify that all variants submitted for publication complied with a widely accepted standard for description. Members of the committee represented expertise in journal publication, variant description, and data sharing. Following a series of monthly committee discussions and consultations with external experts over a 2-year period, a consensus was reached that all published variants should conform to criteria developed by the Human Genome Variation Society (HGVS; <https://varnomen.hgvs.org/>). Furthermore, committee members agreed that authors were responsible for the accuracy of variant description and should, at a minimum, be required to confirm in writing that all variants reported in their manuscripts conformed to HGVS criteria. In addition, the committee recommended that editors request evidence of verification in the form of output files from online publicly available software tools or records documenting submission to online databases that verified nomenclature (e.g. ClinVar [<https://www.ncbi.nlm.nih.gov/clinvar/> , Landrum et al., 2020]; Leiden Open Variation Database [LOVD; <https://www.lovd.nl/> , Fokkema et al., 2011]). Provision of such evidence would assist authors and editors should database curators or others question variant description, as well as identify variants that cannot be defined according to HGVS criteria.

The variant description must adhere to current recommendations of the HGVS (den Dunnen et al., 2016). Variants in the text and tables should be presented at the DNA level, with the option of including (predicted) protein-based designations and RNA-based designations (when RNA is analyzed; see additional considerations below). Ideally, variants will be defined with respect to a specific genomic DNA reference sequence with a reference genome version clearly appended. The latest nomenclature updates, examples of acceptable nomenclature, and guidance concerning reference sequences can be found at <https://varnomen.hgvs.org/> .

Compliance with HGVS nomenclature must be verified using tools such as the Mutalyzer program (<https://mutalyzer.nl/> ; den Dunnen, 2016) or VariantValidator (<https://variantvalidator.org/> ; Freeman et al., 2018) each of which offers a batch mode to facilitate rapid checking of multiple variant descriptions. Alternatively, variants previously submitted to existing online databases that ensure nomenclature compliance (e.g. ClinVar, LOVD), will have been assigned unique IDs by the database in question. Authors should include those unique IDs with the variants in their submitted manuscripts or provide a link retrieving the variant entries from the database.

Authors must provide files generated by verification tools or database submission IDs to confirm that all variants comply with HGVS nomenclature.

Editorial staff will review files provided by authors as a compliance check to confirm that authors used variant validation tools or have submitted their variants to a database that has validated the descriptions.

Authors will be responsible for the accuracy of the variant description in their manuscripts, just as they are responsible for the accuracy of figures, tables, and data files.

Additional important considerations for clear and unambiguous documentation of variant descriptions include: Reference sequences defined in the HGVS nomenclature guidelines (

<http://varnomen.hgvs.org/bg-material/refseq/>) must be used to report sequence variants. Authors should always include the Accession and Version Number of the relevant reference sequence(s) (e.g., RefSeq NM_003002.3 , LRG_9t1, or GenBank NC_000011.10) in the Materials and Methods section (or equivalent) and as a footnote in any tables listing variants.

If alternative nomenclature schemes are commonly found in the literature, they may also be used in addition to approved nomenclature , but they must be defined clearly and unambiguously (e.g., F5 p.Arg534Gln and factor V Leiden).

Standard HGVS nomenclature using “g.” annotation and identifying the genome build must be used for noncoding variants, including those variants identified in genome-wide association studies (GWAS; e.g., NC_000017.11 :g.50201450C>T). Genomic location identifiers from dbSNP (<https://www.ncbi.nlm.nih.gov/snp>) may be added, in addition to approved nomenclature , if the specific nucleotide change is also included.

Among 425 initial submissions to Human Mutation over a 6-month period from April 2019, 349 had variant data and over half (208) included the validation file. During the same time frame, 187 manuscripts under revision reported variants, of which 162 included a validation file, while the remaining 25 indicated that alternative validation software had been used or that variants had been previously reported. Genetics in Medicine (GIM) only checked revised manuscripts for variant data. Over a 3-month period, from April 2019, GIM received 193 revised manuscripts. Thirty-five of these contained variant data and all included a variant file. Four had to be sent back to the authors for errors to be corrected, but all revised manuscripts complied with the requirement to provide a file confirming that validation had been performed before the article was processed as a revision.

Several issues surfaced during the pilot project. Prime among these was the question of whether journal editors or authors should be responsible for correcting nomenclature errors. Two factors contributed to the committee's response to this issue: (1) authors have primary responsibility for the accuracy of data presented in their publications; and (2) editorial staff do not have sufficient expertise or workforce capacity to perform these corrections. Therefore, the committee decided that authors should have final responsibility for correcting the nomenclature of any variants not conforming to HGVS guidelines. In addition, because a substantial fraction of published variants has documented nomenclature errors (e.g., Deans et al., 2016), the committee decided that all variants presented in a manuscript, whether new or previously reported, should be verified. This approach will progressively correct extant nomenclature errors. Mutalyzer and VariantValidator autocorrect variant descriptions and create warnings in their results file, except on rare occasions where there are syntax errors that cannot be interpreted/resolved. The HGVS nomenclature committee can assist when autocorrection does not occur (email to VarNomen@HGVS.org).

Upon reviewing the results of the pilot project, the committee agreed that requiring authors to verify that variants complied with HGVS nomenclature are a reasonable step toward standardizing the worldwide inventory of human variation. The editorial policy can require compliance with the recommendations to describe sequence variants adequately before manuscripts are accepted and published. Ideally, submission of variants and phenotypes to a public database before initial manuscript submission would streamline the review of manuscripts, by virtue of the intrinsic quality checks of compliant public databases (den Dunnen, 2019). Although we encourage immediate public sharing to support rapid use in the diagnosis, some databases (e.g., ClinVar and LOVD) allow an embargo period after data submission to allow for publication time. Finally, the process should not place an undue burden on editorial staff, as it only requires documentary evidence that authors have used the validation tools or have submitted their variants to a database that has validated the descriptions.

Issues that remain for editorial staff to consider for their own journals include the extent to which authors are required to correct nomenclature errors, guidelines for describing variants that cannot be defined using current HGVS description criteria, and manuscript types that may be exempt from description requirements. The committee recognized that a minor fraction of variants has

ambiguous descriptions. Consequently, asking authors to provide documentation that errors could not be resolved after two attempts may be a reasonable compromise. In contrast, a lack of effort on the authors' part to correct remediable errors may be considered as a reason to reject a manuscript.

The HVP/GV Reporting of Sequence Variants Working Group acknowledges that verifying descriptions of DNA variants using validation tools increases workload, albeit minor in most cases, for authors, and for editorial staff. However, we propose that the gain in accuracy outweighs the burden. Similar efforts in the past, such as verification of citations, have improved the robustness of scientific reports. Thus, we encourage all journal editors in human, medical and molecular genetics, and medical journals that frequently report disease-related variants to adopt these minimal recommendations. Table 1 lists the journals which to date have compliant editorial policies. These recommendations are fairly easy to implement, especially in modern editorial offices with automated submission systems and with technical support to editorial office staff from the database administrators. These recommendations will result in a vast improvement of the quality of variant description in the published literature.

Patricia Cornwall for excellent assistance in the preparation of this manuscript; Robert D. Steiner, MD, Editor in Chief, Genetics in Medicine for framing editorial issues and expectations regarding this guidance. Support was provided by the National Human Genome Research Institute (U41HG006834 to Heidi L. Rehm) and the National Institute of Diabetes and Digestive and Kidney Diseases (R01DK044003 to Garry R. Cutting) of the National Institutes of Health.

DISCLOSURES Karen E. Weck: Unpaid conflicts: President, Association for Molecular Pathology.

Heidi L. Rehm, Garry R. Cutting: NIH funding for work related to this manuscript.

Johan T. DenDunnen: Cochair of HGVS nomenclature Sequence Variant Description Working Group (SVD-WG).

Peter J. Freeman: Colead on the VariantValidator project; Member of HGVS nomenclature Sequence Variant Description Working Group (SVD-WG).

Raymond Dalglish: Nonfinancial conflicts: Codeveloper of VariantValidator; Member of the Board of Directors of HGVS; Member of HGVS nomenclature Sequence Variant Description Working Group (SVD-WG).

LIST OF URLs Human Variome Project <https://www.humanvariomeproject.org/>

Human Genome Variation Society <https://varnomen.hgvs.org/>

ClinVar <https://www.ncbi.nlm.nih.gov/clinvar/>

HGVS nomenclature guidelines <http://varnomen.hgvs.org/bg-material/refseq/>

Human Mutation <https://onlinelibrary.wiley.com/journal/10981004>

Genetics in Medicine <https://www.nature.com/gim/>

VariantValidator <https://variantvalidator.org/>

These authors declare no conflict of interests: Jan Higgins, Li Gong, Teri E. Klein, David N. Cooper, Greg Barsh, Adya Misra, Issei Imoto, Peter J. Freeman, Johan T. den Dunnen, Juergen K.V. Reichardt, Bruce Korf, Katsushi Tokunaga, Huw Dorkins, Sarah Ratzel, Sara Cullinan, and Mark H. Paalman.

10.1002/humu.22978 26919400 den Dunnen JT (2016). Sequence variant descriptions: HGVS nomenclature and mutalyzer . *Current Protocols in Human Genetics* , 90 (1), 7.13.1 – 7.13.19 .

10.1002/cphg.2 27367167 den Dunnen JT (2019). Efficient variant data preparation for Human Mutation manuscripts: Variants and phenotypes . *Human Mutation* , 40 (8), 1009 .

10.1002/humu.23830 31397982 den Dunnen JT , Dalglish R , Maglott DR , Hart RK , Greenblatt MS , McGowan-Jordan J , Roux AF , Smith T , Antonarakis SE , & Taschner PEM (2016). HGVS recommendations for the description of sequence variants: 2016 update . *Human Mutation* , 37 (6), 564 – 569 .

10.1002/humu.22981 26931183 Fokkema IFAC , Taschner PEM , Schaafsma GCP , Celli J , Laros JFJ , & den Dunnen JT (2011). LOVD v.2.0: The next generation in gene variant databases . *Human Mutation* , 32 (5), 557 – 563 .

10.1002/humu.21438 21520333 Freeman PJ , Hart RK , Gretton LJ , Brookes AJ , & Dalglish R (2018). VariantValidator: Accurate validation, mapping, and formatting of sequence variation descriptions . *Human Mutation* , 39 (1), 61 – 68 .

10.1002/humu.23348 28967166 PMC5765404 Landrum MJ , Chitipiralla S , Brown GR , Chen C , Gu B , Hart J , Hoffman D , Jang W , Kaur K , Liu C , Lyoshin V , Maddipatla Z , Maiti R , Mitchell J , O'Leary N , Riley GR , Shi W , Zhou G , Schneider V , ... Kattman BL (2020). ClinVar: Improvements to accessing data . *Nucleic Acids Research* , 48 (D1), D835 – D844 .

10.1093/nar/gkz972 31777943 PMC6943040 Tack V , Deans ZC , Wolstenholme N , Patton S , & Dequeker EMC (2016). What's in a name? A coordinated approach toward the correct use of a uniform nomenclature to improve patient reports and databases . *Human Mutation* , 37 (6), 570 – 575 .

10.1002/humu.22975 26920083 Yen JL , Garcia S , Montana A , Harris J , Chervitz S , Morra M , West J , Chen R , & Church DM (2017). A variant by any name: Quantifying annotation discordance across tools and clinical databases . *Genome Medicine* , 9 (1), 1 – 14 .

10.1186/s13073-016-0396-7 28122645 PMC5267466 TABLE 1 Examples of journals that have editorial policies requiring validation of variant nomenclature and their author guidelines/policies