**EMQN Policy on Variant Nomenclature**  
*(v10, updated 2019)*

**BACKGROUND**
Over the years the EMQN has noted wide disparities in the naming of DNA sequence variants. This inconsistency is unacceptable and leads to confusion both in the laboratory and the clinical setting (Tack, Deans, Wolstenholme, Patton, & Dequeker, 2016). To avoid confusion and clinical errors it is important that diagnostic laboratories lead the way in adopting a consistent approach to naming variants.

Nomenclature guidelines are available from the Human Genome Variation Society (HGVS), the international body for defining gene variation nomenclature under the umbrella of the Human Genome Organization (HUGO) and the International Federation of Human Genetics Societies (IFHGS). It is therefore appropriate that we should follow these guidelines and the EMQN recommends that all labs use them (website: [http://varnomen.hgvs.org/](http://varnomen.hgvs.org/)). However we do recognise that there remain some difficulties with the interpretation of these guidelines.

**POLICY**
We have produced the following notes to assist in the interpretation of the nomenclature guidelines.

1. **The EMQN expects laboratories to report their results using HGVS nomenclature** ([http://varnomen.hgvs.org/](http://varnomen.hgvs.org/)) recommending that laboratories implement the latest version (currently v19.01). EMQN will not usually deduct marks for minor HGVS errors and will allow reasonable timescale for laboratories to adapt their reporting policies when new versions of HGVS are published (usually a complete annual cycle of EQA schemes). During this time period comments will be made but no points will be deducted for failure to implement new guidance. EMQN does not expect laboratories to amend their nomenclature for guidance which has not yet been accepted e.g. proposed changes still the subject of community consultation.

   a. When using LRGs it is important to specify the transcript used e.g. COL1A1 LRG_111

   b. If there is no LRG always quote the reference sequence AND the version number as sequences may change from version to version in NCBI/EBI databases.

2. **For all diagnostic purposes, the reference sequence we recommended to be used is the LRG Locus Reference Genomic (LRG: [http://www.lrg-sequence.org/home](http://www.lrg-sequence.org/home)) if one is available. LRGs are more stable and do not contain version numbers. However, RefSeqs from NCBI or EBI files are accepted when no LRG or a ‘provisional’ LRG is available.**

   a. When using LRGs it is important to specify the transcript used e.g. COL1A1 LRG_111

   b. If there is no LRG always quote the reference sequence AND the version number as sequences may change from version to version in NCBI/EBI databases.

3. **Numbering:** Follow the HGVS guidelines and number bases starting with the A of the ATG initiation codon as base 1 (it is also recommended to clearly state that nucleotide numbering starts with the A of the ATG translation initiation site – this is particularly important in genes where legacy numbering does not start from the ATG e.g. BRCA1 and BRCA2).

   a. Legacy numbering is seen in some genes where variants are named from base 1 of a cDNA reference sequence. This can lead to confusion and inappropriate testing e.g. BRCA2 gene variants in the BIC database using accession number U43746.1 where the A of the ATG codon is at position 229.

   b. **Legacy numbering is only allowed in addition to approved numbering.** Any different numbering system to that used in the reference sequence (e.g. constituent in legacy numbering), must be noted on your report (for example; state that the numbering system starts with nucleotide 1 being the beginning of the transcript).

4. **Legacy nomenclature:** It is recognised that some recurrent variants come with historical names which do not comply at all with the HGVS recommendations (e.g. ‘DeltaF508’ in the CFTR gene or ‘Factor V Leiden variant’ in the FV gene). It appears unrealistic that the use of such legacy names in the clinical context will ever be replaced by systematic nomenclature. Therefore, it is recommended that HGVS nomenclature is used in addition to the legacy name when appropriate.

   a. If application of the HGVS guidelines results in a different nomenclature from that used in the past, then you should quote both old and new nomenclatures to avoid confusion with previous reports of the same variant.

5. **In a disease such as cystic fibrosis, laboratory reports frequently list the panel of variants tested. In order to avoid a lengthy list of both old and HGVS compliant nomenclature, only variants identified in the patient or family need to be quoted using HGVS nomenclature.**
6. Where the information is available it is recommended that genetic variants are described at both the DNA and the protein level. Due to redundancy in the genetic code it is not sufficient to describe variants at the protein level only. The protein description is designated as the unprocessed AA chain. If the protein is a theoretical prediction, this should be indicated using parentheses e.g. p.(Arg21Val).

7. If the testing method used does not unambiguously identify a specific base change, the nomenclature used should reflect this. New nomenclature are now available for incompletely specified bases (http://varnomen.hgvs.org/bg-material/standards/#aagcode) which use IUPAP-IUB nucleotide/amino acid codes e.g. B for A,G or T or K for G or T or N for A,G,C or T. Similarly the symbol Xaa may be used for an unspecified or unknown AA in the protein description. Finally the symbol ‘^’ can also be used to indicate ‘or’.

8. In the full disease-specific and molecular pathology EQAs, laboratories are required to use HGVS nomenclature to describe variant alleles, but are not required to provide full HGVS genotypes. For example, it is acceptable to state: “...is heterozygous for the CFTR c.1521_1523delCTT p.(Phe508del) variant”, rather than stating the full genotype c.[1521_1523delCTT];[=] p.([Phe508del]);[=].. Additionally it is acceptable to describe an allele with no clinically relevant variants in words rather than using HGVS e.g. no pathogenic variant identified (homozygous).

9. In the technical EQA scheme for Sanger Sequencing it is no longer required that variants are reported using full HGVS genotype nomenclature e.g. it is acceptable to state: “...is heterozygous for the CFTR c.1521_1523delCTT p.(Phe508del) variant”, rather than stating the full genotype c.[1521_1523delCTT];[=] p.([Phe508del]);[=].. It is acceptable to describe an allele with no clinically relevant variants in words rather than using HGVS e.g. no pathogenic variant identified (homozygous).

10. Measurement of zygosity in tumour DNA is problematic: samples may have normal cell contamination, multiple cell lines, or amplifications and it is therefore acceptable to omit zygosity for tumour samples. Please see section 11 below.

11. It is recommended that laboratories follow HGVS guidance on the naming of mosaic variants:
   a. c.85C=/>T - The sample is a mix of cells containing c.85C= and c.85C>T

12. Laboratories reporting the presence of exon deletions and duplications will not be penalised for not using HGVS nomenclature (but HGVS may also be used). A clear statement of which exons are deleted/duplicated is acceptable. An appropriate reference sequence must be given which clearly conveys the numbering of the exons to the reader. For example, a Locus Reference Genomic (LRG) sequence plus transcript number could be used if published, as this includes numbered exons. If an NCBI RefSeq has been used to derive the numbering you should indicate if the exon numbering is systematic or custom (legacy numbering). As with other tests, the methodology must be stated in the report and if a commercial kit used then the kit name and version number must be provided.
   a. Please note changes to HGVS MLPA nomenclature are currently still the subject of community consultation and are therefore not currently recommended by EMQN.

13. Laboratories reporting the presence of triplet repeat variants, or other large scale genomic changes (for example, results of methylation analysis, uniparental disomy, copy number analysis, uniparental disomy analysis ) will not be penalised for not using HGVS nomenclature. However, where a specific gene [rather than genomic region] is being tested then, to properly identify the gene, the RefSeq (LRG if available) and the HGNC gene symbol for that gene should be included in all reports.

14. Please note if laboratories report full genotypes then they must be given using correct HGVS nomenclature. For 2019, if minor errors regarding the placement of brackets, semi-colons etc. are observed then no marks will be deducted but feedback comments will be given. If major errors are observed such as incorrect nucleotide or amino acid then appropriate marks will be deducted in the genotyping category.

15. Describing the normal allele: Recent changes to HGVS nomenclature have led to a variety of different descriptions of the normal allele depending on the situation e.g. a predictive test, sequencing of the whole cDNA, sequencing of a single exon etc. To avoid unnecessarily penalising laboratories EMQN will accept a written description e.g. no pathogenic variant identified, heterozygous for..., or an appropriate HGVS annotation e.g. c.456C=

**FEEDBACK**

Please let us know of any difficulties in applying these guidelines. Feedback from you will help us establish consensus recommendations.

**REFERENCES**


**SUMMARY OF CHANGES TO DOCUMENT**
<table>
<thead>
<tr>
<th>Version no.</th>
<th>Change(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>New document</td>
</tr>
<tr>
<td>2</td>
<td>Updated email address</td>
</tr>
<tr>
<td>3</td>
<td>Updated address</td>
</tr>
<tr>
<td>4</td>
<td>Major rewrite - Statement added that labs should use v2 nomenclature.</td>
</tr>
<tr>
<td>5</td>
<td>Updated to include new sections on use of full HGVS, duplication/deletions and other large scale genomic changes. Also clarified background and policy and added a section for feedback. Indicated version and update date on title so that labs see when reviewed.</td>
</tr>
<tr>
<td>6</td>
<td>Updated policy with recommendations on tumour DNA - not using full HGVS and how to describe mosaicism if you do not know if the change is somatic or germline</td>
</tr>
<tr>
<td>7</td>
<td>Title changed to “Variant” Nomenclature from “Mutation” nomenclature</td>
</tr>
<tr>
<td>8</td>
<td>Major rewrite to address HGVS version 15.11</td>
</tr>
<tr>
<td>9</td>
<td>Alternatives for naming the normal allele following introduction of HGVS v15.11</td>
</tr>
<tr>
<td>10</td>
<td>Policy 9: Sanger scheme no longer requires the use of ‘full’ HGVS to specify zygosity, Policy 14 updated to position on minor HGVS errors, and Document title: HGVS version updated to 19.01</td>
</tr>
</tbody>
</table>