20140904_data_release_notes

Overview of changes in the release of September 4, 2014

Overview of submissions: 2014

Date	Total Submissions
Jan 01, 2014	68204
Feb 01, 2014	73492
Mar 01, 2014	83343
Apr 01, 2014	111501
May 01, 2014	112349
Jun 01, 2014	117209
Jul 01, 2014	127132
Aug 01, 2014	127557
Sep 1, 2014	143114

Content

Brief	Explanation
Clinical significance for somatic variants submitted by OMIM	Based on personal communication from Ada Hamosh, ClinVar altered the method of calculating clinical significance from the title of the allelic variant record. If the word 'SOMATIC' is present, the value for clinical significance is now calculated as 'Pathogenic' instead of 'other'. (http://www.ncbi.nlm.nih.gov/clinvar/docs/clinsig/). This altered the interpretation of approximately 500 variations.
Intron variant as molecular consequence	This release includes more comprehensive reporting of intronic variants, cross referenced to the sequence ontology value of SO:0001627. The term was added to variants within a single intron (i.e. excluding those that span one or more exons) that are not already encoded as occurring in splice donor or acceptor sites. The term used to describe these was also changed from intron to intron variant.

AllHighlyPenetrant becomes 'not specified'	Based on a recommendation from the ClinGen group, the use of AllHighlyPenetrant as the name of a phenotype was phased out. AllHighlyPenetrant had been used to represent a set of disorders for which a pathogenic allele would be expected to be expressed as that disorder. It was created for use in ClinVar so that submitters could convey the concept that an assertion of *benign* clinical significance for an allele applied to more than one disorder. ClinVar will now use 'not specified' when *benign* or *not provided* is submitted as clinical significance and no specific disorder is reported. On the web display, no value will be shown when 'not specified' is in the underlying data.
Findings and "See cases"	The XML is now being more consistent with MedGen in classifying traits by type. Observed clinical features, rather than diagnostic terms, are now represented under ObservedIn/TraitSet in the XML, with @type ="Finding", rather than a set of traits about which an assertion is being made with @Type ="Disease". When observed clinical features are reported, without a diagnostic term, the phase 'See cases' is reported in //ClinVarAssertion/TraitSet/Trait/Name. This modifications affect primarily the submissions from ISCA (http://www.ncbi.nlm.nih.gov/clinvar/submitters/500029/)
unknown not uncertain	When the origin of a variant is not clear, NCBI has decided to use the term 'unknown' rather than 'uncertain'. ClinVar records, and the XSD, were updated accordingly.
SequenceLocation for duplications	Representation of the sequence location of short tandem duplications had been reported inconsistently for the last few months in the SequenceLocation element. Namely, some had been treated as insertions, between nucleotides where an insertion could have occurred, and some had been reported according to the location of the duplicated sequence. Records have been updated to be consistent with the latter convention. This convention remains under review.
Submissions	This release includes both a comprehensive update and new submissions from the International Standards for Cytogenomic Arrays (ISCA) Consortium.

Overview of Submissions: 2013

Date	Total Submissions
Apr 05, 2013	30333
May 01, 2013	30386
Jun 01, 2013	39047
Jul 01, 2013	39170
Aug 01, 2013	45901
Sep 01, 2013	50263
Oct 01, 2013	52047
Nov 01, 2013	64750

Dec 01, 2013	64881