Deliverable 3.1
OD4RD Report on Genetic Annotations (new and updated gene-disease relationships), ORPHAcodes alignments with other terminologies, and newly produced or updated texts (definitions and abstracts)

March 2023

Co-funded by the European Union

Co-Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or HADEA. Neither the European Union nor the granting authority can be held responsible for them

Content summary

Introduction

The main scope of OD4RD Work Package 3 (WP3) was to improve the Orphanet scientific information collected around the Nomenclature of rare diseases, either as a continuous internal updating activity or in collaboration with European Reference Networks, to reflect the evolution of knowledge.

This summary report will provide an overview of the main activities from January 2022 to March 2023 and it will contain 3 sections:

- New discoveries around genes related to RD and their representation in the Orphanet database, with links to where these relationships are available for reuse;
- RD coverage in target terminologies and the relationships between ORPHAcodes to each of the target terminologies, with links to where these alignments are available for reuse;
- A list of newly produced definitions and of newly produced or updated RD summary information, with information on whether the text was realized in collaboration with ERNs.

New discoveries around genes related to RD

Methodology

Updates of the gene-diseases relationships database are regularly performed by the Orphanet gene project manager according to a methodology published here (link). Updates in the gene-disease database can also result in a revision of the Orphanet nomenclature (addition or modification of an ORPHAcode) and follows the methodology for updating the Orphanet nomenclature. If these updates impact the revision done in collaboration with ERNs, the appropriate connections are done to avoid parallel revisions.

Data report

The following indicators were calculated on the time period going from January 2022 to March 2023.

Cumulative number of genes linked to rare diseases with a disease-causing relationship	**4,462**
Newly created disease-causing genes (i.e. genes were absent from Orphanet and were created with a disease-causing relationship)	**61**
New disease-causing relationships attributed to unknowingly disease-causing genes (i.e. genes were present in Orphanet as parts of diagnostic panels and were subsequently identified as disease-causing)	**53**
Newly added disease-causing relationships attributed to knowingly disease-causing genes (i.e. genes already present in Orphanet linked to other diseases as disease-causing)	**130**
Suppressed gene-disease relationships	**119**
Modification of gene-disease relationships (e.g. disease-causing germline mutation was updated to disease-causing germline mutation loss of function)	**7**

The complete Orphanet dataset of genes associated with rare diseases and their relationships is generated every 6 months (July and November) and available for download and exploitation in the Orphadata website, Orphanet Scientific Knowledge Base section (link).

RD coverage in target terminologies

Methodology

Alignment of ORPHAcodes with other terminologies in use in health systems and in registries is regularly performed by the Orphanet alignments project manager under the supervision of the Orphanet Scientific Director. The objectives of the OD4RD project were axed on four different terminologies, with different advancement statuses:

- Complete and maintain the alignment between ORPHAcodes and ICD-10 codes according to the rules already established and published here (link). This process implies the identification of the best suitable ICD-10 code and the definition of a semantic relationship between both codes;
- Continue and deliver for the first time the alignments between ORPHAcodes and ICD-11, in the frame of the participation of Inserm, US14-Orphanet to the French WHO collaborating centre. Two complementary approaches are followed to complete the alignments and to identify the gaps in ICD-11 in terms of rare disease coverage: a) a syntactic approach, using language recognition algorithms performed by the Agence Numérique en Santé (ANS, lead of the WHO collaborating centre). Results of the application of these algorithms are curated by Orphanet and combined with b) a semantic approach, derived from the same rules used for ICD-10 alignments and needed to assess the proximity (exact or inexact) of both Orphanet and ICD-11 concepts.
- Continue the mapping exercise between ORPHAcodes and SNOMED-CT in the frame of the collaboration agreement in place between Orphanet and SNOMED International, in order to have a common release in October each year and containing new terms. This collaboration implies monthly follow-up meetings and exchange of information, in order to reach agreement on the concepts to be included in SNOMED-CT and mapped to ORPHAcodes. Orphanet releases a human-readable excel file with the updated mappings every year, and for the duration of the collaboration agreement.
- Continue the alignments between ORPHAcodes and OMIM concepts by capturing OMIM updates by running a script on a monthly basis, then manually curating the concepts in order to produce semantic relationships between concepts in both terminologies or update existing ones. This activity can also result in initiating a demand for updating the Orphanet nomenclature, and follows the methodology for updating the Orphanet nomenclature. Should these updates impact the revision done in collaboration with ERNs, the appropriate connections are done to avoid parallel revision.

Data report

The following indicators were calculated on the time period going from January 2022 to March 2023.

ORPHAcodes aligned to ICD-10:	Total:	7,288	
	Disorders:	6,171	(98,4% aligned)
	Subtypes of disorders:	1,003	(98,3% aligned)
	Groups of disorders:	114	(5,34% aligned) *
ORPHAcodes aligned to ICD-11:	Total:	1,711	
	Disorders:	1,285	(20,5% aligned)
	Subtypes of disorders:	108	(10,5% aligned)
	Groups of disorders:	318	(14,9% aligned) *
ORPHAcodes aligned to SNOMED-CT:	Total:	6,397	
	Disorders:	5,868	(93,5% aligned)
	Subtypes of disorders:	301	(29,5% aligned)
	Groups of disorders:	228	(10,7% aligned) *
ORPHAcodes (genetic entities only) aligned to OMIM:	Total:	4,526	
	Disorders:	3,915	(86,3% aligned)
	Subtypes of disorders:	559	(72,2% aligned)
	Groups of disorders:	52	(3,90% aligned) *

Table 1: ORPHAcodes aligned to target terminologies. Percentages are calculated on the total of Orphanet active clinical entities, by classification level, for ICD-10, ICD-11 and SNOMED-CT, while they are calculated on the total of Orphanet active genetic clinical entities for OMIM. As of March 2022, Orphanet total number of clinical entities is 9424, of which 6270 are disorders, 1020 are subtypes of disorders and 2134 are Groups of disorders. Orphanet total number of genetic clinical entities is 6639, of which 4533 are disorders, 774 are subtypes of disorders and 1332 are Groups of disorders.

*Please note that for group of disorders, Orphanet only aligns the exact concepts in all terminologies

The complete Orphanet dataset of RD coverage in target terminologies and the relationships between ORPHAcodes to each of the target terminologies is generated every 6 months (July and November, except for SNOMED-CT, that is released annually in October) and is available for download and exploitation in the Orphadata website, Orphanet Scientific Knowledge Base section (link).

Newly produced definitions and newly produced or updated RD summary information

Methodology

Production or update of textual information related to rare disorders either as a continuous internal updating activity or in collaboration with European Reference Networks. This activity follows a general methodology published here (link).
Collaborations can be established with ERNs in order to prioritize the common work to maintain up-to-date summary information on rare diseases in Orphanet, as well as the definitions that accompany every rare disorder and that are delivered together with the nomenclature to help assigning the right ORPHAcode to a patient.

The collaboration with ERNs follows these steps:

- Initiation of collaboration: there are three points of entry 1) following the finalization of a collaboration on the Orphanet nomenclature 2) The ERN establishes contact for a large collaboration independently of a nomenclature collaboration 3) an ERN is approached directly by the Orphanet Editorial team to update a selection of diseases;
- Identification of training required: for large collaborations an initial introduction to Orphanet along with an Editorial training is useful. For smaller collaborations, experts are directly issued guidelines on the task and text content;
- Establishing a list of diseases and text type (definition quality control or production/update of a disease abstract) according to specialty and predefined set of priorities;
- Identification of suitable experts by the ERN coordinating team or working group leads. Diseases involving multiple specialties may require cross-ERN collaboration;
- Cycle of revision and curation of text(s) between expert(s) and Editorial team, resulting in the validation and publication of text(s);
- Maintenance of disease summary information (typically every 3 years);
- Review and feedback of collaboration: given punctually throughout collaboration and at the end of a collaboration cycle.

This methodology has been established based on the previous ERN collaboration experience and is adaptable to the needs and resources of each network.

Data report

The following indicators were calculated for the time period going from January 2022 to March 2023.

A total of 238 texts (either abstracts or quality control definitions) have been produced, modified (either minor or major modification) and/or validated by an expert reviewer. Of those a total of 128 texts were reviewed by an expert reviewer and published, 103 of those published texts were reviewed by an expert affiliated to an ERN. 110 texts are still in a pre-published phase, and of those 26 texts are being reviewed by an expert affiliated to an ERN.

ERN	N. of published texts	N. of ongoing texts
Endo-ERN	18	4
EpiCARE	1	/
ERKNet	6	/
ERN CRANIO	1	/
ERN EURO-NMD	5	3
ERN GENTURIS	1	/
ERN ITHACA	7 (one cosigned with eUROGEN)	1
ERN RARE-LIVER	/	1
ERN RITA	/	/
ERN-EYE	/	1
ERNICA	5	/
ERN-LUNG	3	/
ERN-RND	/	2
ERN-Skin	25	3
EURACAN	8	/
EuroBloodNet	10	/
eUROGEN	5 (one cosigned with ERN ITHACA)	/
MetabERN	9	11

A total of 175 definitions have also been internally produced and 71 of those have been published following internal medical validation. The remaining definitions are still in a pre-published phase and awaiting medical validation.

The complete list of newly produced or updated RD summary information (with a focus on those realized in collaboration with ERNs) and of newly produced definitions is available in the Annex .