Orphan Drug Development Guidebook Task Force

Galaxy Guide for drug development

Building Block FACT SHEET FORM content

This document defines the content of the FACT SHEET to be created for each identified tool, incentives, initiative or practice (the Building Block) introduced by public bodies or used by developers to expedite drug development in Rare Diseases (RDs).

ITEM	DESCRIPTION
Building Block (BB) Title	Orphanet database
References	www.orpha.net www.orphadata.org Activity report: https://www.orpha.net/orphacom/cahiers/docs/GB/ActivityReport2017.pdf
Description	Established in 1997, Orphanet is a unique public resource worldwide, gathering and improving knowledge on all rare diseases (RD), affecting less than 1 in 2'000 people in the European population. Orphanet derives from its multi-lingual, manually curated and expert validated knowledge base an ontology of RD (Orphanet Rare Disease Ontology, ORDO), information on RD (www.orpha.net) and data (www.orphadata.org). Orphanet maintains the Orphanet RD nomenclature (see Codification BB), improving the visibility of RD in health and research information systems, acting as an interoperability vector between healthcare and research (https://drive.google.com/file/d/1dsfAAdxF4USmsKK9YQ4IpDW44OqNIZ--/view?usp=sharing). The nomenclature is semantically aligned with: OMIM, ICD, SNOMED-CT, MedDRA, UMLS, MeSH, GARD. RD are annotated with manually curated data on age of onset, age of death, prevalence, incidence, gene-disease qualified relationship, cross-references with other databases (OMIM, UniProtKB, HGNC, ensembl, Reactome, IUPHAR, Genatlas), the frequency of phenotypic features (using HPO), and International Classification of Functioning (ICF-CY) derived terms. Orphanet knowledge base includes, besides the nomenclature, scientific annotations (phenotypes, disabilities, epidemiological data, genes, age of onset and of death), textual information (produced by Orphanet or from tother sources after quality assessment), orphan designations and drug data from EMA and FDA, and a catalogue of expert services/resources in Orphanet consortium countries described with meta-

ITEM	DESCRIPTION
	data, including: expert centres, diagnostic tests and labs, patient orgnisations, patient registries, biobanks, research projects, clinical trials and research infrastructures. Orphanet, and its nomenclature is now an internationally recognized standard: International Rare Disease Research Consortium (IRDiRC) awarded Orphanet and ORDO IRDiRC Recognized Resource status in 2015. These two resources were also named Human Variome Project Recommended Systems in 2017.
Relevance to rare disease drug development	As far as the second IRDiRC goal is concerned, namely 1000 new therapies for rare diseases to be approved by 2027, the majority of which will focus on diseases without approved options, efficient translation from basic and pre-clinical research to clinical trials and therapies discovery is required: "this important goal can be achieved only through a dramatically more efficient development process driven by a radically new approach utilising common standards across distinct research fields, sharing of best practices, creating sustainable business models, and redefining the regulatory environment." Orphanet and its ontological representation, ORDO, provides a common language between the healthcare and research field, improving interoperability between electronic health records, codification systems, registries and cohorts, variant databases, and biobanks. In Europe, European Reference Networks have adopted the ORPHA nomenclature for sharing data through their eHealth platform and for their registries. In addition, Orphanet database allows for analysis of the research landscape in Europe (i.e. identification of trends and gaps) and for identification of where the resources are (i.e.: patient registries, data repositories…).
Category	Regulatory BB
Availability	Data is RD specific and available for all types of publics at the following conditions: • Products promoting the interoperabity of rare disease data on www.orphadata.org: free, open access, CC BY 4.0 1. Orphanet nomenclature and definitions, classifications and cross-references 2. Disease-gene annotations and cross-references 3. Disease-phenotype annotations 4. Epidemiological/natural history annotations 5. Orphanet Rare Disease Ontology and HPO-ORDO Ontological Module (HOOM) • Data sets requiring DTA for academia/ fee for industry: 1. Textual information 2. Catalogue of expert services/ressources 3. Orphan drug data
Geographical scope	Orphanet database is based at the Inserm, in France. Coverage of expert resources include 37 countries

ITEM	DESCRIPTION
	https://www.orpha.net/orphacom/cahiers/docs/GB/Orphanet_Network_MB_members.pdf
Scope of use	Academic researchers and the pharmaceutical industry use Orphadata datasets for research and development purposes. Examples of pharma industry use cases include the incorporation of Orphadata datasets in pre-competitive tools in order to prioritise therapies development, as well as epidemiological data to assess market size. Other use cases include datamining technology applications to generate hypotheses from Orphanet data and textual information.
Subject	Academic researchers, SMEs, pharma industry, policy decision makers
Enablers/ Requirements	Any particular requirement for open access datasets (see above) that are available in: • XML • JSON • Sparql EndPoint • Ontology: OWL, txt, obo (and for ORDO : http://bioportal.bioontology.org/ontologies/ORDO (OWL, CSV, RDF/XML) DTA signature for restricted access datasets (academia, public not-for profit institutions); contract (for fee) for private for-profit companies.
Output	**A public website,** www.orpha.net**:** around 30 million pages viewed in 2017, >40,000 daily visitors from 232 countries per day **A download platform (reusable datasets)** www.orphadata.org**.** Orphadata products were downloaded more than 212,000 times, with an average of 17,690 times per month in 2017. **Ontologies:** ORDO (downloaded 7,137 times in 2017) and HOOM, launched in 2017 (available at orphadata.org)
Actors and Stakeholders	Funders: Public: Inserm, French Ministry of Health, European Commission, Agence de la Biomédecine, Ministries of health, universities and hospitals in the countries of the Orphanet network; Private: AFM-Téléthon, Fondation Groupama pour la Santé Pharma companies and bio-techs using for-fee data Academic researchers, public administrations and Patient organisations: using the open access datasets Health professionals, patients, researchers, students, decision makers (Orphanet website users)

ITEM	DESCRIPTION
Use	Epidemiological data: determination of the size of the market per geographical region Recruitment, collaborations: identification of experts, patient organisations, registries, biobanks, research groups developing i.e. animal models, biomarkers, targets, etc, by disease/group of diseases. Pre-competitive tools: integration with other sources of data, including private pharma data, to generate hypothesis.
PROs/ advantages	Computable structured data, manually curated, organized hierarchically (thus allowing for clustering and aggregation of data) and described with unique identifiers. Numerous cross-references with other relevant resources (gene function, pathways, compounds and targets,…) Versioning and differentials for open access data. Documentation provided. A number channels are available to help users wishing to access and re-use Orphanet data. Firstly, there is a contact form on the Orphadata website (http://www.orphadata.org/cgi-bin/contact.php) as well as a dedicated alias data.orphanet@inserm.fr, with a 24 hour first response time during office days. A FAQ and user guide is also available. In addition, there is an ORDO user mailing list (ordo-users.orphanet) to deliver updates concerning the ontology.
CONs/ risks	Dataset formats are standardized and on-demand, customized datasets as a service are not yet developed (but will be) APIs in development, not yet available. Hands-on assistance can be necessary to make the best use of the resource (as proven by the Orphanet participation to the training courses organized by ISS for RD registries: BYOD (RD-Connect/Excelerate). Not already available as a service, but in discussion.
Best time to apply	N/A (Comment: the answer depends on the purpose for which the database content is used; we can provide more explanations if needed)
Duration	DTA or for-fee contracts are one-year in duration, re-conductible. Free, open-access data can be accessed at any time and are monthly updated in

ITEM	DESCRIPTION
	Orphadata.
Cost	Open datasets are for free. For fee datasets (for for-profit organisations) are here: http://www.orphadata.org/cgi-bin/img/PDF/Catalogue_Orphadata_2018.pdf (The catalog will be updated in 2019, in particular: epidemiological data will become open access)
Practical tips	This can include DOs and DON'Ts and strategic considerations.

Ana Rath 9-12-18 20:02
Opmerking [1]: Need clarification to answer this item.