Rare diseases codification in health information systems:

current situation and evolutions forseen

Ana Rath, Annie Olry, Bertrand Bellet, Ségolène Aymé

ORPHANET - Inserm US14, Paris, France

ana.rath@inserm.fr

European Expert Group on Health Information

Luxembourg 26 November 2014

Improved codification for rare diseases is cited as a priority in the Council Recommendation on an action in the field of rare diseases (2009)

Need for coding RD

- **Make RD visible** in order to:
  - Have sound epidemiological data
  - Document the natural history of RD
  - Identify patients from health records for clinical research
  - Bring clinical data to research
- Different systems are using **different terminologies**
  - **Need for inter-operability**
- Need **to have a common language** to allow for sharing clinical data between health care centres and databases and registries:
  - Patients are rare and scattered
  - Significant amounts of data are necessary to perform research

Steps so far

- **2009-2012 RDTF Joint Action**: support cross-referencing terminologies & ICD11 revision process for RD
- **2012 on: EUCERD Joint Action**: continuation of the RDTF work
  - Sept 2012: EUCERD JA workshop on cross-referencing terminologies
  - March 2014: EUCERD JA workshop on Orphacodes in HIS
- **July 2014: CEGRD:** draft recommendation on codification for rare diseases
- **October 2014 (JRC-ISPRA):** Workshop on the next steps to promote the implementation of Orphacodes in MS
- **November 2014**: Adoption of CEGRD recommendation

The current situation: EU MS

Country	Current coding systems used	Interest in/plans to use Orphacode	Plans to implement other system	Approach favoured
Austria	ICD 10	yes		Electronic link
Belgium	ICD 9	yes	Snomed, ICD 10	Parallel coding
Bulgaria	ICD	yes		Parallel coding
Croatia	ICD10	yes		Parallel coding
Cyprus	ICD10 CY			
Czech Republic				
Denmark	ICD10 DM, Nordic codes, OMIM	no		
Estonia				
Finland				
France	ICD10, Orphacode	yes		Parallel coding
Germany	ICD 10 GM	yes		Electronic link
Greece	ICD10	yes (registry)		
Hungary	ICD10	yes		Parallel coding
Ireland	ICD10	yes		

The current situation: EU MS

Country	Current coding systems used	Interest in/plans to use Orphacode	Plans to implement other system	Approach favoured
Italy	ICD9 CM (hospital) ICD10 (statistics) ICF, Orphacodes	yes		Electronic link
Latvia	ICD10	yes (registries)	OMIM	Parallel coding
Lithuania	ICD10	yes (CE)		Parallel coding
Luxembourg	ICD 10	yes		
Malta				
Netherlands	ICD10	yes (CE)		
Poland	ICD10	yes (CE)		Parallel coding
Portugal				
Romania	ICD10, some use of Orphacodes (CE)			
Slovakia	ICD10, OMIM	yes (registries)		Parallel coding
Slovenia	ICD10, OMIM, Orphacode (CE)	yes		
Spain	ICD9 ICD10	no	Snomed, ICD 10	
Sweden	ICD10	no	Snomed, ICD 10	
United Kingdom	ICD10, Snomed			

EUCERD
European Union Committee

EGRD, July 2014

The current situation: terminologies

- Most health information systems use **ICD**
  - Some ICD-9
  - Most ICD-10
- WHO' ICD-11 revision is expected for 2017
- Some countries have adopted **SNOMED CT**
- Genetic databases use **OMIM**

- In Europe, countries having national plans/strategies for RD decided to integrate the Orphanet nomenclature of RD, and code patients with the **ORPHA** code

Terminologies currently used

- **SNOMED CT** (Systematized Nomenclature of Medicine – clinical terms, IHTSDO); 401,200 terms
  - comprehensive clinical terminology,
  - multihierarchical ontology
  - intended for use in EHR, and to semantic interpretation of EHR
  - translated in licensed countries
- **ICD-10** (International classification of diseases- WHO), 12,451 terms
  - Monohierarchical classification of diseases
  - Intended for statistical uses (morbidity, mortality)
  - Translated and adapted in different countries

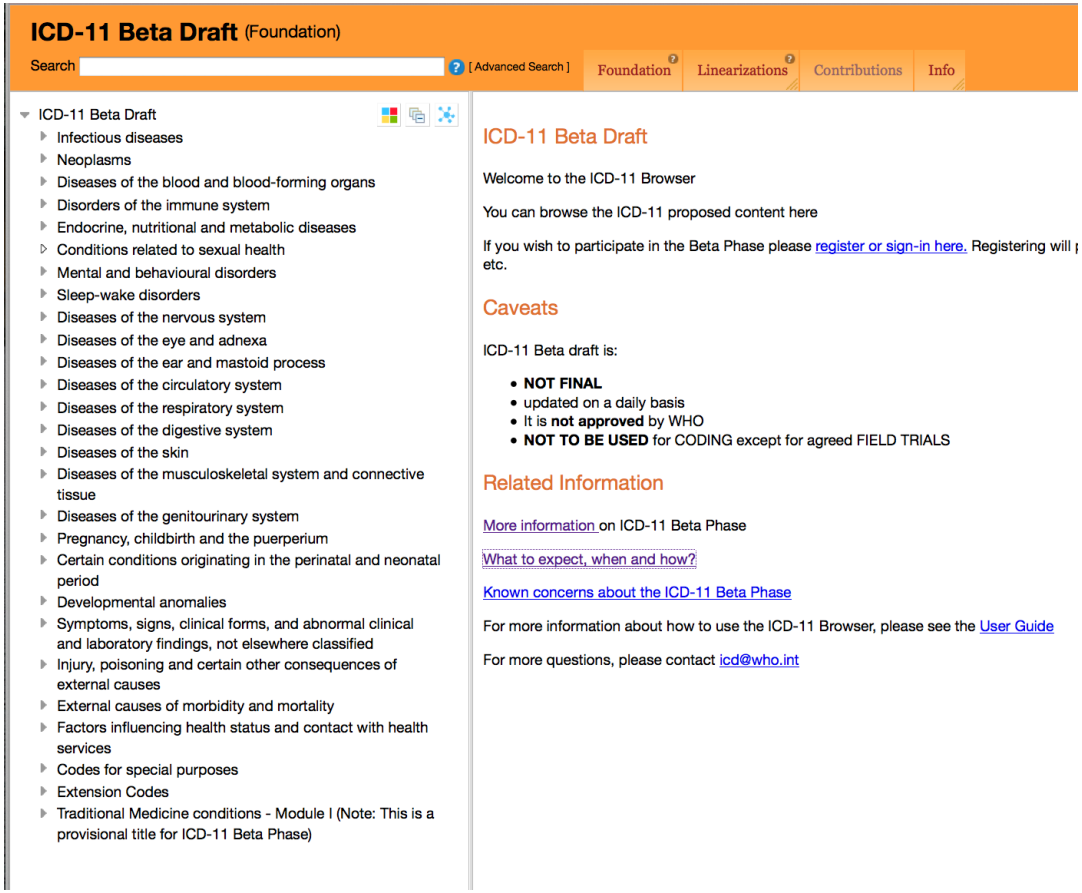
Other terminologies/resources

- **OMIM** (Online Mendelian Inheritance in Man):
  - Genetic disorders and phenotypes (regardless their rarity)
  - Organised by genes
  - English only
  - Use in (genetic) databases

How many RD are included in these terminologies?

- **ICD10**
  - **466 specific codes** matching Orphanet rare disease entities (including groups of diseases) (= EXACT mappings)
  - 431 inclusion terms matching Orphanet RD entities
  - 82 index terms matching Orphanet RD entities
  - → Total: only 979 Orphanet RD entities with an ICD-10 mention
  - But >80% of ORPHA entries have been attributed an ICD10 code
- **SNOMED CT** (from UMLS AA2013)
  - On 15,043 candidate mappings, 3,541 were EXACT (**2,883 ORPHA** entries)
- **OMIM**
  - On 6,617 total mappings, 3,388 are EXACT (**3,380 ORPHA** entries)

The ICD - 11



RD-TAG input since 2009

> 5,000 RD incorporated

> 3,000 definitions
Introduced by Orphanet

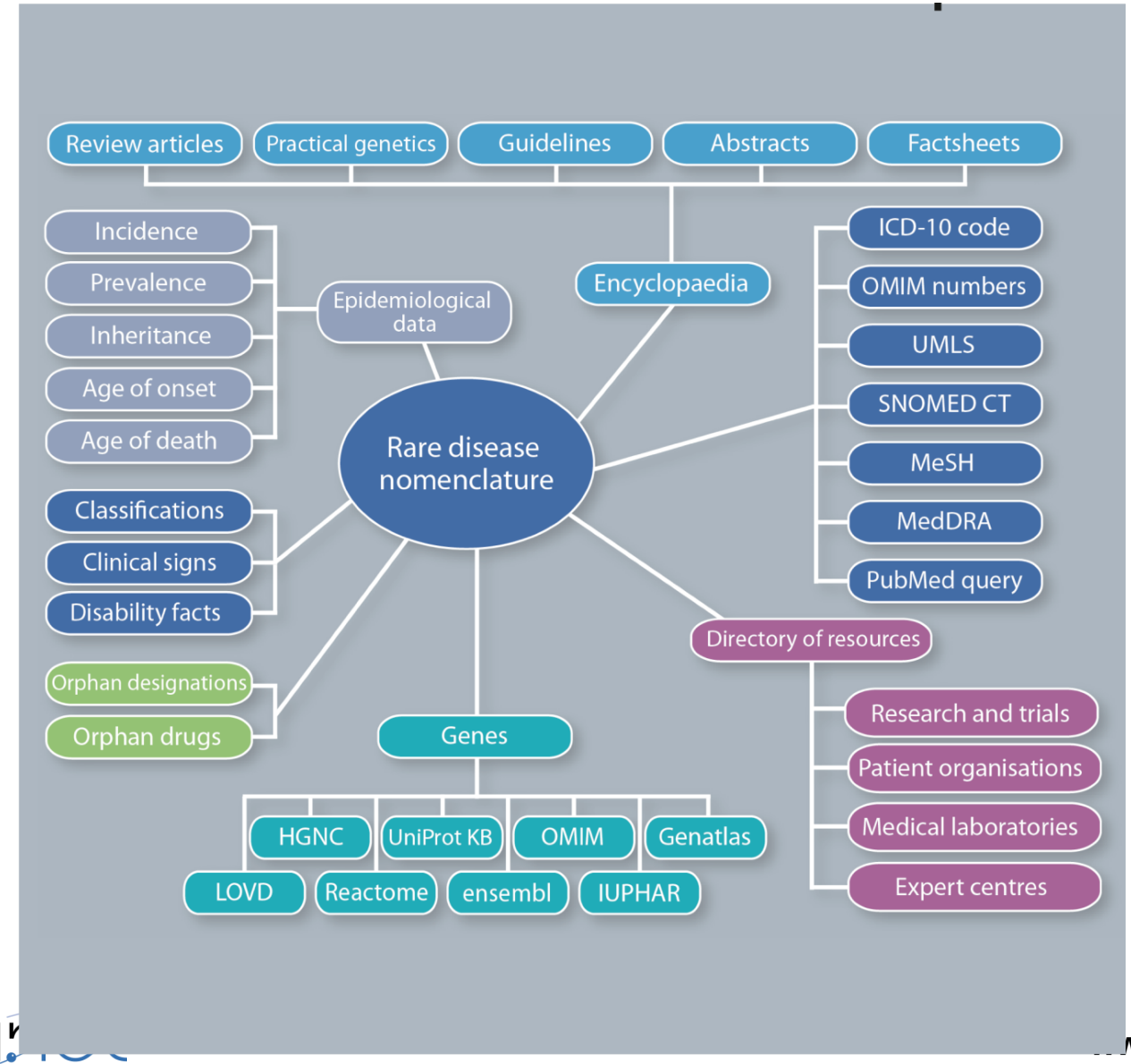
Currently, resolution of known problems, linearizations, tests

Still some conflicts to solve between TAGs

Release expected end 2017

Need for an update process
(ORPHA – ICD11 alignments
To be incorporated in ICD11)

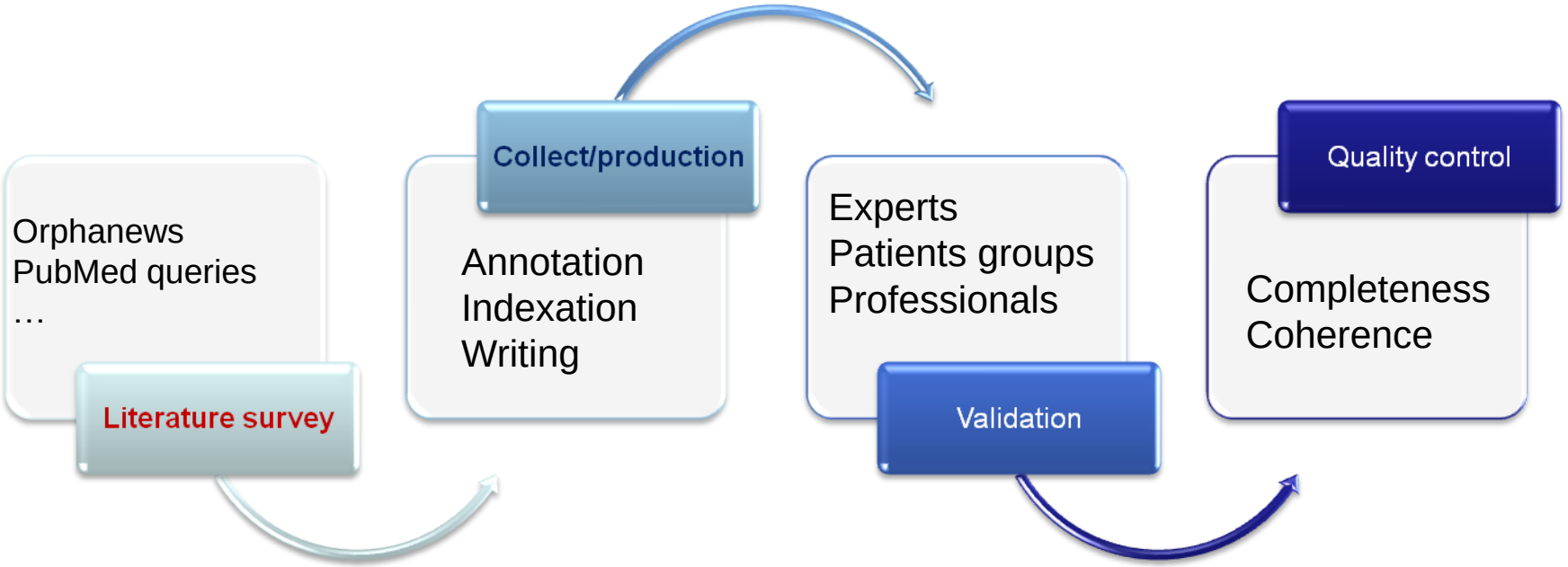
Orphanet , the European database for RD



Rare disorders in Orphanet

- Since 1997: **Inventory** of rare disorders (prev<1/2 000)
  - Mapped to OMIM
- 2005: Mapping to ICD-10
- 2007: **Classification** of rare disorders
- 2011: Mappings to UMLS, SNOMED CT, MeSH, MedDRA
- 2014: ORDO (Orphanet **ontology** of rare diseases) in collaboration with the EBI.

Edition process



Orphanet nomenclature

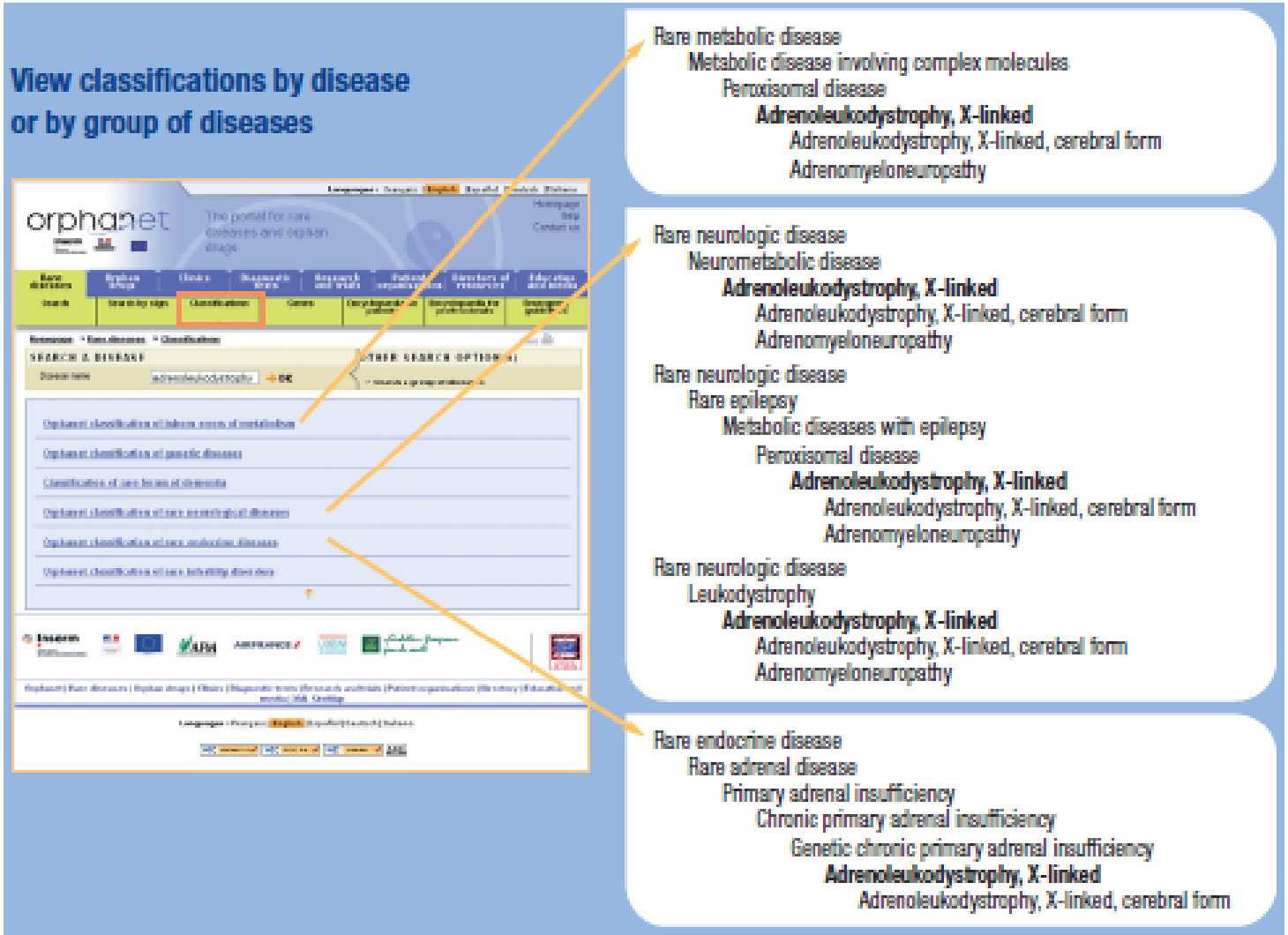
- Each entry (group, disorder, subtype) is given a unique, stable **ORPHA number**
  - The Orphacode
- Each entry is given a preferred term and as many synonyms as necessary
- Nomenclature is translated (FR – ES – IT – PT – DE – NL)
- Nomenclature and classifications are updated monthly.

Orphanet nomenclature

- Produced by the INSERM (France), hosting Orphanet, the EU RD database
  - Institutional support
- Under CreativeCommons
  - Free for re-use
- Strong position in international terminologies
  - Ongoing process to be a WHO collaborating center
    - Population and updates of ICD11
  - Ongoing MoU with IHTSDO (SNOMED CT)
    - Population of SNOMED CT and mappings

A multidimensional classification

- Orphanet classifications by medical specialty based on international literature and experts advice



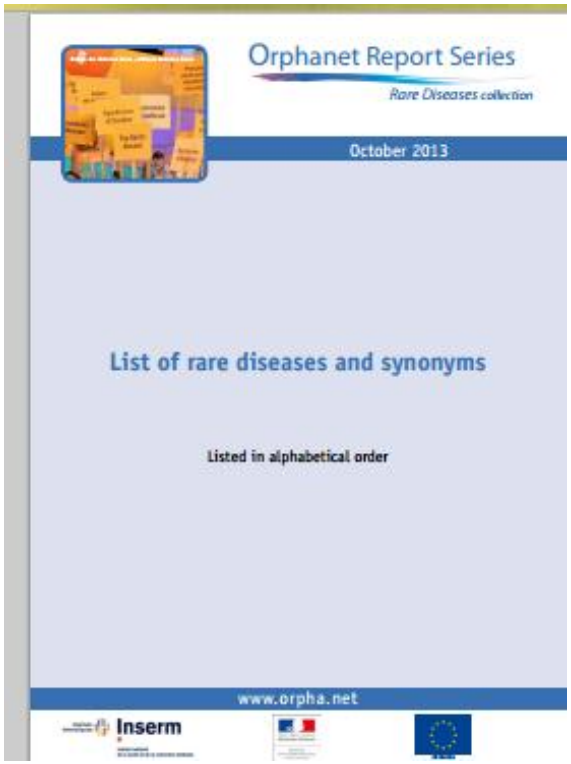
Mappings to other terminologies

- Disorders mapped to OMIM (manually)
- Disorders mapped to ICD-10 (manually)
- Disorders mapped to UMLS, MeSH, SNOMED CT, MedDRA (semi-automatically)
- Mappings are qualified (exact ; narrow-to-broad ; broad-to-narrow)
- Information on the validation status is noted
- Updates depending on the target terminology
  - Monthly (ICD10, OMIM)
  - Twice a year (UMLS, SNOMED CT, MeSH, MedDRA)

Qualifying mappings

E	exact mapping (the terms and the concepts are equivalent)
NTBT	narrower term maps to a broader term
BTNT	broader term maps to a narrower term
W	incorrect mapping (two different concepts)
NTBT/E	narrower term maps to a broader term because of an exact mapping with a synonym in the target terminology
BTNT/E	broader term maps to a narrower term because of an exact mapping with a synonym in the target terminology
W/E	incorrect mapping (two different concepts) but syntactically exact mapping to a synonym or a preferred term in the target terminology
ND	not yet decided/unable to decide
The following are attributed to ICD10 codes only :	
Specific code	The term has its own code in the ICD10
Inclusion term	The term is included under a ICD10 category and has not its own code
Index term	The term is oncluded in ICD10 index and refers to one more general code
Attributed code	The term does not exist in ICD10 and a code was attributed by Orphanet

To find the Orphanet nomenclature



Alphabetical list of names and synonyms
with ORPHA code
Updated every 6 months
Translated.
PDF (download, print)

ORPHA nomenclature
Cross-references
Monthly updated
6 languages (7 soon)
For download
XML

Free access data from Orphanet
orphanet

orphadata

Home

About Orphadata

About Orphanet

Access Orphanet[→]

Contact

Freely accessible datasets

Diseases, cross referenced with other nomenclatures

February-2014

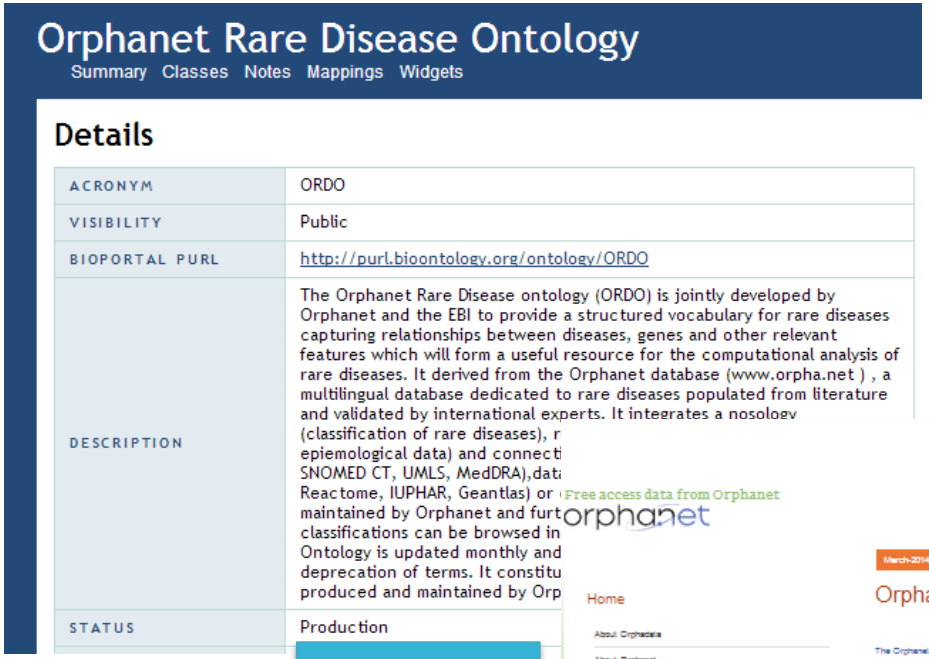
Rare Diseases And Cross-Referencing

Files available in XML format.

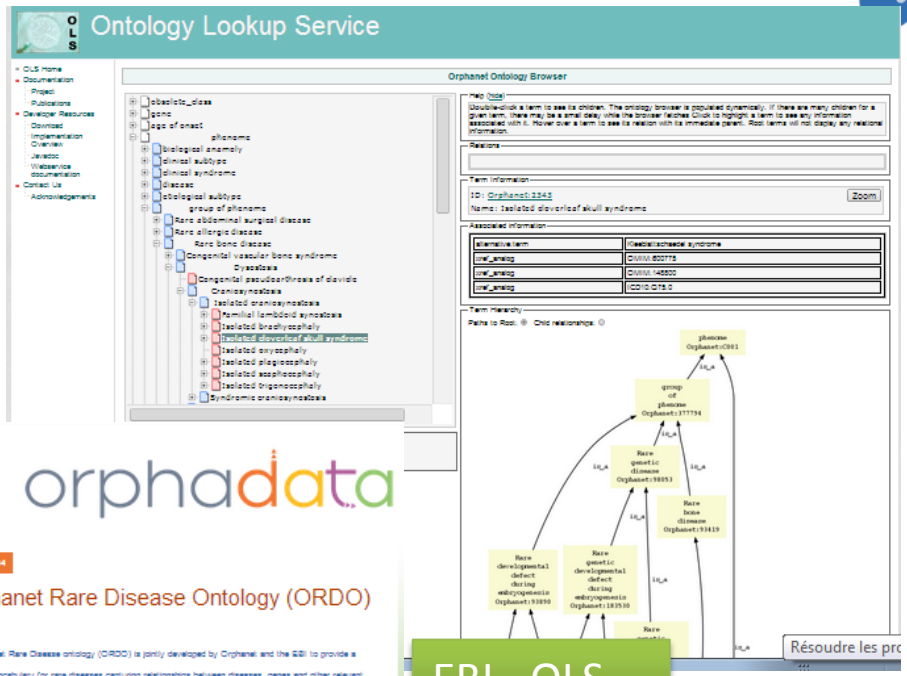
Rare diseases and cross-referencing

Language	links	size
English	http://www.orphadata.org/data/xml/en_product1.xml	6.14 MB
French	http://www.orphadata.org/data/xml/fr_product1.xml	6.15 MB
Spanish	http://www.orphadata.org/data/xml/es_product1.xml	6.10 MB
Italian	http://www.orphadata.org/data/xml/it_product1.xml	6.05 MB
Portuguese	http://www.orphadata.org/data/xml/pt_product1.xml	5.81 MB
German	http://www.orphadata.org/data/xml/de_product1.xml	5.99 MB

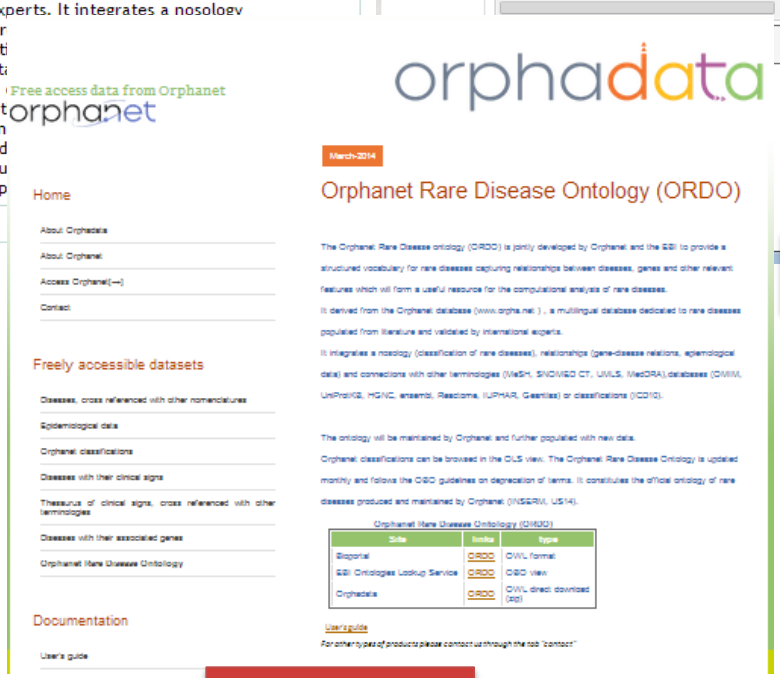
ORDO



BioPortal



EBI - OLS



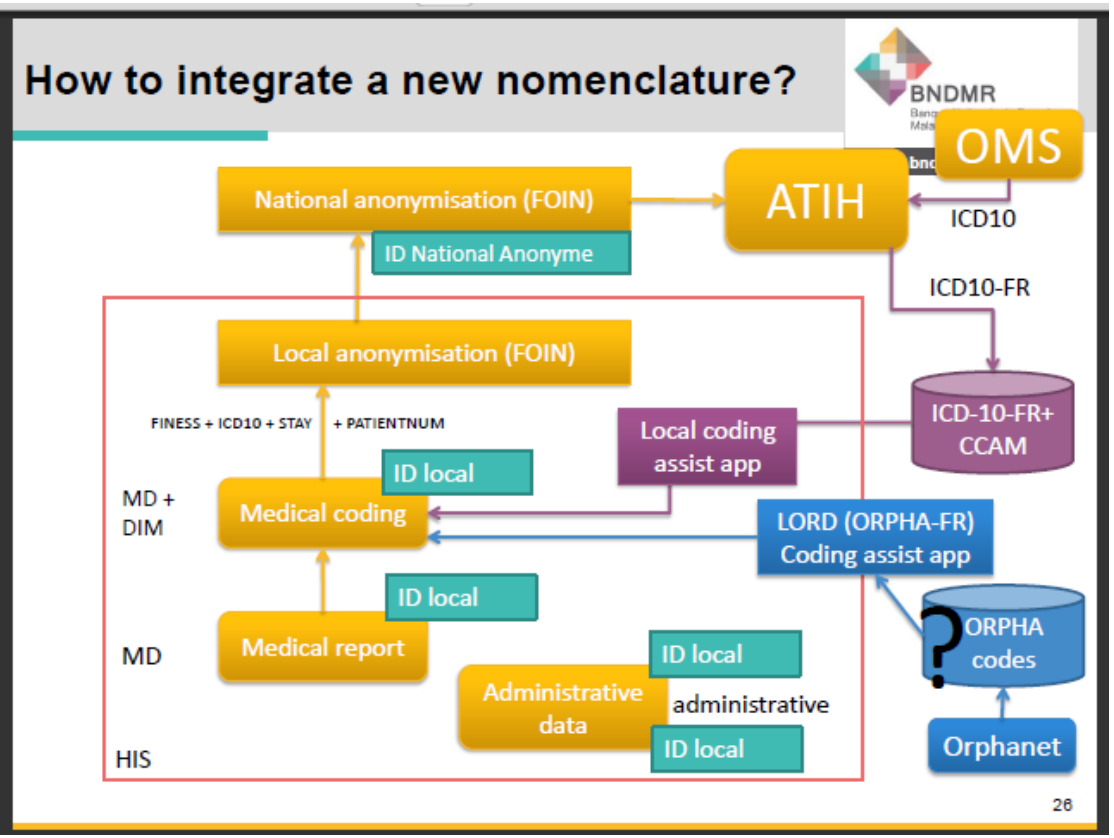
Orphadata

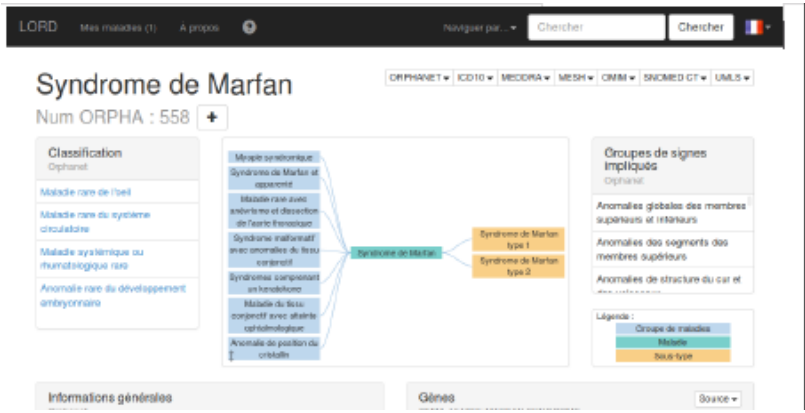
orphanet

www.orpha.net

Ongoing initiatives implementing Orphacodes

France: ORPHA besides ICD10





Lord.bndmr.fr

Rémy Choquet, BNDMR project leader, ISPRA, 2014

Germany: Orpha aligned to ICD10

ICD10-GM / alpha-ID

ORPHA code assigned to

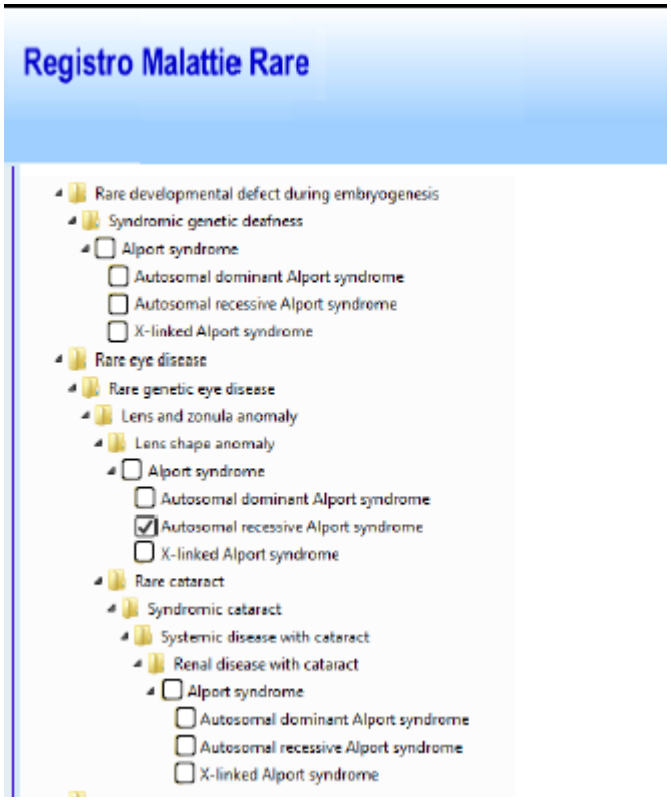
ICD10-GM / alpha-ID

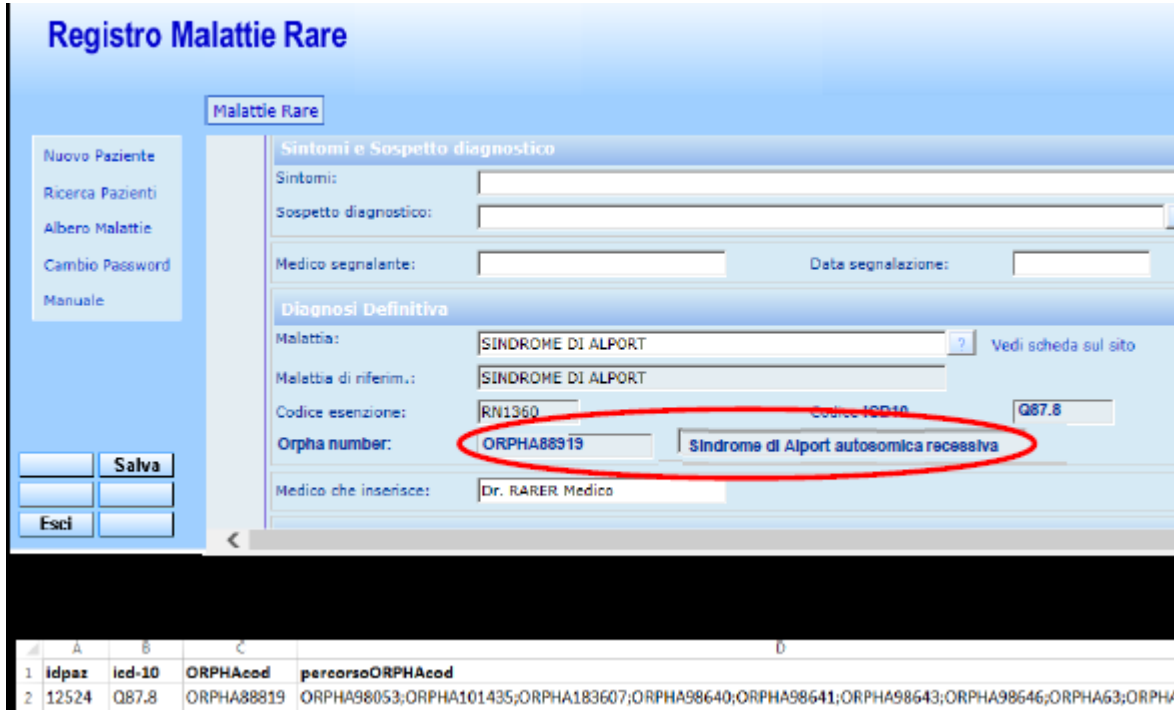
Alpha-ID	ICD-10-GM code	Asterisk code	Orpha number	Term
I15988	Q77.0		932	Achondrogenesis
I17761	Q79.3		2368	Gastroschisis
I83662	Q87.2		2614	Nail-patella syndrome
I82099	Q87.2		2614	Turner-Kieser syndrome
I24937	K90.8+	M14.8*	3452	Whipple disease
I75193	H35.0		40923	Eales disease
I9770	H53.5		49382	Achromatopsia

Stephanie Weber, DIMDI, ISPRA 2014

Italy: alignments + multi-hierarchies

Automatic tool
Based on oRPHA – ICD10 alignments
Retrieving information on the health care utilisation





Paola Facchin, Venento registry, ISPRA 2014

Needs at MS and EU level

- 195 Orphanet classification of rare immunological diseases (Pat)
  - 13022 Rare immune disease
    - 14933 Primary immunodeficiency
      - 18068 Primary immunodeficiency due to a defect in adaptive immunity
        - 14908 Combined T and B cell immunodeficiency
          - 18303 Severe combined immunodeficiency
            - 21578 T-B+ severe combined immunodeficiency
              - 170 T-B+ severe combined immunodeficiency due to gamma chain deficiency
              - 10366 T-B+ severe combined immunodeficiency due to JAK3 deficiency
              - 17829 T-B+ severe combined immunodeficiency due to IL-7Ralpha deficiency
              - 17830 T-B+ severe combined immunodeficiency due to CD45 deficiency
              - 17831 T-B+ severe combined immunodeficiency due to CD3delta/CD3epsilon/CD3zeta
              - 19062 Severe combined immunodeficiency due to CORO1A deficiency
            - 21579 T-B- severe combined immunodeficiency
              - 993 Severe combined immunodeficiency due to DCLRE1C deficiency
              - 1258 Short-limb skeletal dysplasia with severe combined immunodeficiency
              - 8023 Severe combined immunodeficiency due to adenosine deaminase deficiency
              - 10318 Reticular dysgenesis
              - 20431 Severe combined immunodeficiency due to LCK deficiency
              - 21580 Severe combined immunodeficiency due to DNA-PKcs deficiency
              - 21997 Severe combined immunodeficiency due to complete RAG1/2 deficiency
            - 22703 T+ B+ severe combined immunodeficiency
              - 22219 Severe combined immunodeficiency due to CARD11 deficiency
                - Genes
                  - 22042 CARD11 (Disease-causing germline mutation(s) in)
              - 22702 Severe combined immunodeficiency due to IKK2 deficiency
                - Genes
                  - 22826 IKBKB (Disease-causing germline mutation(s) in)

Needs at EU level

Definition of a common level of interoperability
&
Codification standards

Next steps

- WP in the Rare diseases joint action 2015-2017
- Steering group
  - Common denominator at EU level
  - Guidelines and coding instructions
  - Definition of a common resource
    - To ensure consistency in quality of data
    - To allow interoperability between MS

Thank you !