

orphanet

---

orphadata

**USER'S GUIDE**

---

June 2011

## TABLE OF CONTENTS

<b>RARE DISEASES AND CROSS-REFERENCES.....</b>	<b>2</b>
<b>EPIDEMIOLOGICAL DATA.....</b>	<b>4</b>
<b>ORPHANET CLASSIFICATIONS.....</b>	<b>6</b>
<b>CLINICAL SIGNS .....</b>	<b>8</b>
<b>THESAURUS OF CLINICAL SIGNS.....</b>	<b>10</b>
<b>DISEASES WITH THEIR ASSOCIATED GENES .....</b>	<b>11</b>

## PART I

---

### RARE DISEASES AND CROSS-REFERENCES

This file contains the inventory of rare diseases listed in the Orphanet database. Each disease is described by its main name and synonyms. It is also defined by its ORPHA number, and mapped with ICD-10, the 10<sup>th</sup> International Classification of Diseases established by the World Health Organization (<http://www.who.int/classifications/icd/en/>), and OMIM, the Online Mendelian Inheritance in Man database (<http://www.omim.org/>).

This file is available in 6 different languages: English, French, Italian, Spanish, German and Portuguese. If terms are not yet translated from English, they remain in English by default.

The threshold applied to define a disease as rare is the one defined in Europe: any disease or condition with a prevalence in the general population equal to or lower than 5 in 10,000.

#### Description of the XML content

- **DiseaseList count:** total number of rare diseases in the XML file.
- **Orphanum:** unique identifying number assigned by Orphanet to a given disease.
- **ExpertLink:** stable URL pointing to the specific page of a given disease on the Orphanet website.
- **Lang:** ISO 639 code for language names.
- **Name:** preferred name of a given disease.
- **SynonymList:** synonyms for a given disease name.
- **ExternalReferenceList:** list of cross-references for a given disease in the ICD-10 and OMIM systems.
- **Source:** can be either OMIM or ICD-10.
- **Reference:** listed reference for a given source associated with a disease.

## Example

**<DiseaseList count="6234">**

6234 is the total number of diseases presented in this XML file

**<Orphanum>58</Orphanum>**

The ORPHA number of this disease is 58

**<ExpertLink lang="en">http://www.orpha.net/consor/cgi-bin/OC\_Exp.php?lng=en&Expert=58</ExpertLink>**

The stable URL pointing to information on this disease is

[http://www.orpha.net/consor/cgi-bin/OC\\_Exp.php?lng=en&Expert=58](http://www.orpha.net/consor/cgi-bin/OC_Exp.php?lng=en&Expert=58)

**<Name lang="en">Alpha-mannosidosis</Name>**

The main name of the disease in English is "Alpha-mannosidosis"

**<SynonymList count="1"> <Synonym lang="en">Lysosomal alpha-D-mannosidase deficiency</Synonym>**

This disease name has one synonym "Lysosomal alpha-D-mannosidase deficiency"

**<Source>OMIM</Source> <Reference>208400</Reference>**

**<Source>ICD10</Source> <Reference>E77.1</Reference>**

This disease is defined by two external references: one MIM number (208400) and the ICD-10 code E77.1

## PART II

---

### EPIDEMIOLOGICAL DATA

This file contains epidemiological data for rare diseases. Each disease is associated with its class of prevalence, type of inheritance, average age of onset and average age of death.

This file is available in 6 different languages: English, French, Italian, Spanish, German and Portuguese. If terms are not yet translated from English, they remain in English by default.

The data are extracted from the literature as to reflect the situation in Europe. The validity of the published studies is taken for granted and not re-assessed, although there is a low level of consistency between studies and usually poor documentation of methods used.

#### Description of the XML content

- **DiseaseList count:** total number of rare diseases in the XML file.
- **Orphanum:** unique identifying number assigned by Orphanet to a given disease.
- **ExpertLink** stable URL pointing to the specific page of a given disease on the Orphanet website.
- **Lang:** ISO 639 code for language names.
- **Name:** preferred name of a given disease.
- **ClassOfPrevalence:** estimated prevalence of a given disease. There are eight possible values: >1 / 1,000, 1-5 / 10,000, 6-9 / 10,000, 1-9 / 100,000, 1-9 / 1,000,000 or <1 / 1,000,000, No data available, Unknown.
- **AverageAgeOfOnset:** estimated average age of disease onset. There are six different population age groups: Neonatal/infancy, Childhood, Adolescence / Young adulthood, Adulthood, Variable, No data available.
- **AverageAgeOfDeath:** estimated average age at death for a given disease. There are eight different population age groups: Before age 5, Child / adolescent, Young adult, Adult, Elderly, Any age, Normal, No data available.
- **TypeOfInheritance:** type(s) of inheritance associated with a given disease. There are ten different types of inheritance: Autosomal dominant, Autosomal recessive, X-linked dominant, X-linked recessive, Chromosomal inheritance, Mitochondrial inheritance, Multigenic/multifactorial, Sporadic, No data available, Unknown.

## Example

**<DiseaseList count="6234">**

6234 is the total number of diseases presented in this XML file

**<Orphanum>61</Orphanum>**

The ORPHA number of this disease is 61

**<ExpertLink lang="en">http://www.orpha.net/consor/cgi-bin/OC\_Exp.php?lng=en&Expert=61</ExpertLink>**

The stable URL pointing to information on this disease is

[http://www.orpha.net/consor/cgi-bin/OC\\_Exp.php?lng=en&Expert=61](http://www.orpha.net/consor/cgi-bin/OC_Exp.php?lng=en&Expert=61)

**<Name lang="en">Alpha-mannosidosis</Name>**

The main name of the disease in English is "Alpha-mannosidosis"

**<Name lang="en">1-9 / 1 000 000</Name> </ClassOfPrevalence>**

Estimated disease prevalence is between 1 / 1,000,000 and 9 / 1,000,000

**<Name lang="en">Neonatal/infancy</Name> </AverageAgeOfOnset>**

Average age of disease onset is during the neonatal/infancy period

**<Name lang="en">Any age</Name> </AverageAgeOfDeath>**

Average age of death can be at any age

**<Name lang="en">Autosomal recessive</Name> </TypeOfInheritance>**

The disease is inherited as an autosomal recessive trait

## PART III

---

### ORPHANET CLASSIFICATIONS

This file contains the hierarchical representation, *i.e.* classification, of medical domains. There is one file per classification.

This file is available in English only.

Rare diseases in Orphanet, depending on their clinical presentation, are included in as many classifications as needed. Classifications are based on published scientific articles and reviewed by experts.

#### Description of the XML content

- **Orphanum:** unique identifying number assigned by Orphanet to a given classification or disease.
- **ExpertLink:** stable URL pointing to the specific page of a given disease or classification on the Orphanet website.
- **Lang:** ISO 639 code for language names.
- **Name:** preferred name of a given disease / group of diseases or classification.
- **ClassificationNodeList count:** number of diseases or group of diseases present at the same level of the classification.
- **ClassificationNodeChildList count:** number of diseases or group of diseases classified under a given disease / group of diseases.

## Example

**<Orphanum>68334</Orphanum>**

The ORPHA number of this group of diseases is 68334

**<ExpertLink lang="en">http://www.orpha.net/consor/cgi-bin/OC\_Exp.php?lng=en&Expert=68334</ExpertLink>**

The stable URL pointing to information on this group of diseases is  
http://www.orpha.net/consor/cgi-bin/OC\_Exp.php?lng=en&Expert=68334

**<Name lang="en">Rare hemorrhagic disorder due to a constitutional coagulation factors defect</Name>**

The name of this group of diseases in English is "Rare hemorrhagic disorder due to a constitutional coagulation factors defect"

**<ClassificationNodeChildList count="2">**

Two diseases are present at this classification level

**<Orphanum>162949</Orphanum>**

The classification ORPHA number is 162949

**<Name lang="en">Orphanet classification of rare hematological diseases</Name>**

This is the name of the classification

**<ClassificationNodeChildList count="15">**

There are 15 diseases classified under the group of "Rare hemorrhagic disorder due to a constitutional coagulation factors defect"



## PART IV

---

### CLINICAL SIGNS

This file contains rare diseases listed in Orphanet with their associated clinical signs and symptoms, using a controlled vocabulary (Thesaurus) organised by main organs and systems.

As the indexing process is still ongoing, this information is only currently available for about 2,000 rare diseases.

This file is available in 6 different languages: English, French, Italian, Spanish, German and Portuguese. If terms are not yet translated from English, they remain in English by default.

#### Description of the XML content

- **DiseaseList count:** total number of rare diseases in the XML file.
- **Orphanum:** unique identifying number assigned by Orphanet to a given disease.
- **Lang:** ISO 639 code for language names.
- **Name:** preferred name of a given disease or a clinical sign.
- **DiseaseSignList count:** number of clinical signs associated with a given disease.
- **Sign:** a given clinical sign.
- **SignFreq:** estimated frequency of occurrence for a given clinical sign in a given disease. Three different frequency groups have been defined: occasional, frequent and very frequent.

## Example

**<DiseaseList count="6234">**

6234 is the total number of diseases presented in this XML file

**<Name lang="en">Alexander disease</Name>**

The main name of the disease is "Alexander disease"

**<DiseaseSignList count="6">**

There are 6 clinical signs associated with this disease

**<DiseaseSign>**

**<Name lang="en">Macrocephaly/macrocrania/megalocephaly/megacephaly</Name>**

**<Name lang="en">Very frequent</Name> </SignFreq>**

Macrocephaly/macrocrania/megalocephaly/megacephaly is a clinical sign very frequently seen in patients with Alexander disease.

## PART V

### THESAURUS OF CLINICAL SIGNS

This file contains the thesaurus of clinical signs used for indexing rare diseases. Clinical signs are defined using a controlled vocabulary and classified by main organs and systems.

This file is available in 6 different languages: English, French, Italian, Spanish, German and Portuguese. If terms are not yet translated from English, they remain in English by default.

#### Description of the .TXT file content

- **PrtTag:** direct upper-level group of clinical signs to which a given clinical sign belongs.
- **SgnTag:** clinical sign name in English.
- **Lng:** ISO 639 code for language name.
- **SgnLbl:** translation of a clinical sign name from English to a given other language.

#### Example

PrtTag:	SgnTag:	Lng:	SgnLbl:
	Skull/cranial anomalies	de	Schädel-/ kraniale Anomalien
Skull/cranial anomalies	Microcephaly	de	Mikrozephalie

Microcephaly ("Mikrozephalie" in German) is a clinical sign belonging to the group of Skull/cranial anomalies ("Schädel-/ kraniale Anomalien" in German)

## PART VI

---

### DISEASES WITH THEIR ASSOCIATED GENES

This product contains the inventory of rare diseases listed in the Orphanet database and the associated genes. Each gene is described by its official name and synonyms, defined by the HUGO Gene Nomenclature Committee (HGNC). Genes are mapped with their *HGNC ID*, *OMIM number*, *GenAtlas* symbol and *UniProtKB/Swiss-Prot* references.

*HGNC* (<http://www.genenames.org/>) is a committee jointly funded by the US National Human Genome Research Institute (NHGRI) and the Wellcome Trust (UK). It operates under the auspices of HUGO, with key policy advice from an International Advisory Committee. It is in charge of approving gene names and symbols (short-form abbreviations). All approved symbols are stored in the HGNC database. Each symbol is unique and each gene is only given one approved gene symbol.

*OMIM*, Online Mendelian Inheritance in Man (<http://www.omim.org/>), is the database of human genes and genetic phenotypes.

*GenAtlas* (<http://www.genatlas.org/>) is a database of genes and phenotypes. Only the objects with a known cytogenetic location are retained.

*UniProtKB* (<http://www.uniprot.org/>) is the central hub for the collection of functional information on proteins, with accurate, consistent and rich annotation.

This product is available in English only.

#### Description of the XML content

- **DiseaseList count:** total number of rare diseases in the XML file.
- **Orphanum:** unique identifying number assigned by Orphanet to a given disease or gene.
- **Name:** preferred name of a given disease or a gene.
- **GeneList count:** number of genes associated with a given disease.
- **Symbol:** official HGNC-approved gene symbol.
- **Synonym list:** list of synonyms for a given gene.

- **External Reference List:** list of references in HGNC, OMIM, GenAtlas and UniProtKB associated with a given gene.
- **Source:** HGNC, OMIM, GenAtlas or UniProtKB.
- **Reference:** listed reference for a given source associated with a gene.

## Example

**<Orphanum>61</Orphanum>**The ORPHA number of this disease is 61

**<Name lang="en">Alpha-mannosidosis</Name>**

The main name of the disease in English is "Alpha-mannosidosis"

**<GeneList count="1">**

The disease is associated with one gene

**<Orphanum>123328</Orphanum>**

The gene ORPHA number is 123328

**<Symbol>MAN2B1</Symbol>**

**<Name lang="en">Mannosidase, alpha, class 2B, member 1</Name>**

Its official symbol and name are MAN2B1 and Mannosidase, alpha, class 2B, member 1, respectively

**<SynonymList count="2">**

**<Synonym lang="en">LAMAN</Synonym>**

**<Synonym lang="en">MANB</Synonym>**

There are two synonyms for this gene: LAMAN and MANB

**<ExternalReferenceList count="4">**

This gene is mapped with 4 other references

**<Source>GENATLAS</Source>**

**<Reference>MAN2B1</Reference>**

The reference for this gene in GenAtlas is MAN2B1

<Source>**HGNC**</Source>

<Reference>**6826**</Reference>

The reference for this gene in HGNC is 6826

<Source>**OMIM**</Source>

<Reference>**609458**</Reference>

The reference for this gene in OMIM is 609458

<Source>**UNIPROTKB/SWISSPROT**</Source>

<Reference>**O00754**</Reference>

The reference for this protein in UniProtKB is O00754