RD-Action WP5
–
Specification and implementation manual of the Master file for statistical reporting with Orphacodes

Second Part of Milestone 27:
"A beta master file version to be tested in some selected MS together with the correct coding procedures"

Part of Deliverable 5.3:
"An European integrated master file"

Co-funded by
the Health Programme
of the European Union

This Manual is part of the project / joint action ‘677024 / RD-ACTION’ which has received funding from the European Union’s Health Programme (2014-2020).
It has been produced by the co-leaders of the Work Package 5 and is part of the Task 5.2: Specification of the required resources for coding RD consistently across Europe (Task Leader: Dr. Stefanie Weber [DIMDI, Germany] - Contributors: All WP5 contributors). It constitutes the second part of Milestone 27: “A beta master file version to be tested in some selected MS together with the correct coding procedures” and one part of Deliverable 5.3: “An European integrated master file”.

The RD-ACTION Joint Action was launched in June 2015 for a 36 months period.
More information on the activities of the RD-ACTION Joint Action can be found at www.rd-action.eu

Table of Content

1. Introduction

The Specification and implementation manual of the Master file for statistical reporting with Orphacodes (MF) has been developed in the frame of the Rare Disease Joint Action Workpackage 5 (WP5) “Steering, maintaining and promoting the adoption of Orphacodes across Member States” (MS). It shall help to understand and work with the MF. With the Specification and implementation manual, it will be possible for the WP5 participating countries to test the MF, get support on the alignment of the pre-existing national coding system with the Orphanet classifications and give a feedback to the WP5-partners (contact details at the end of the document). According to the countries feedback the MF will be further developed and adapted to the needs of MS. On a long-term view, the MF enables coding of rare diseases (RD) with a standardized code-set for national and international reporting for statistics as well as promoting an enhanced semantic interoperability between country specific IT systems and registries. It can also be used to support a standardized coding of RD patients on national and international level.

The benefits of using this manual as well as the concept of the MF and how to work with it are depicted. Additional to that, the guidelines, the release schedule and further information on the topic are described in the following chapters.

2. Benefits of using the Master file for statistical reporting with Orphacodes

As already mentioned RD-Action WP5 started to develop a file in order to support RD data sharing across MS. Through the use of such file, namely the MF, the process of collection of coded RD-data should be simplified and structured. The MF should facilitate the standardised use of the Orphanet Nomenclature by providing only the data that is relevant for data sharing at EU-level and thus enable international statistical aggregation. Offering an alignment of diagnostic terms with the Orpha number and a terminal ICD-10 code in a dataset, should minimize the bureaucratic burden of using different classifications and support standardization. Furthermore, giving the MS the possibility to align such datasets to other classifications used locally for coding could give a higher added value to the file. Nevertheless, the MF allows an implementation at different levels so that all MS could use it independent of the available resources. (Please see chapter 4 of this document and Deliverable 5.2 for further explanation)

Integrating the International Classification of Diseases (ICD) and the Orphanet nomenclature into one file simplifies the process of assigning an Orphacode to a RD-patient. The coder does not have to select the codes from two separate coding systems but can select them simultaneously by selecting the relevant disease name as the MF combines both.

If the Orphacodes are integrated in a national coding system through other means, it should be ascertained that statistical reporting of coded RD-data will comply with the granularity and content of the data collected through integration of the MF into a national coding system. As the MF is compiled by using the Orphanet-classifications, the same mechanism can be used to aggregate data with more granularity to the level of Orpha numbers contained in sheet two of the Masterfile.

The linkage of the Orphacodes and the ICD codes (or other national coding systems) increases as well the standardization of rare diseases coding and consequently increases interoperability on data sharing. The interoperability of data is a key aspect in processes where data sharing plays a role. In the rare diseases field data sharing becomes much more relevant than in common diseases. On one hand, there is a higher need of networking between health care providers given the reduced number of experts in each disease. And on the other hand, there is the need to aggregate data for epidemiological purposes and research on international level given the reduced number of patients.

At the end, the MF and the guidelines (see chapter 7) should enable data sharing at EU-level and ensure consistency in codification with Orphacodes amongst member states.

3. Concept of the Master file for statistical reporting with Orphacodes

To serve all benefits described above, it is necessary to provide a clear concept of the MF.

The MF is a subset of the Orphanet classifications of RD that contains only the Orpha numbers that should be used for data sharing and statistical purposes at EU-level. Its final version will be produced in an annual basis and will be published by Orphanet, as the so called MF. (For details on how it is produced, see chapter 5).

The list of Orpha numbers included in the MF can of course be part of a more complex setting but specifies the minimum agreeable level for European reporting and comparison. Besides the implementation of the MF, users are free to use the whole Orphanet nomenclature to collect data in a more detailed way, if desired. This means that the MF could be implemented in the diverse data collection systems that use the Orphanet nomenclature. Emphasis should be placed on the possibility to automatically retrieve data at any time according to the Orphanumber list in the MF and according to the basic coding rules and guidelines given by WP5 in Deliverable 5.2. Depending on the way the Orpha codification has been locally implemented, Orphanet classifications may be needed to reach the common level of granularity.

4. Content definition of the Master file for statistical reporting with Orphacodes

With the help of the Orphanet Team the WP5 developed an algorithm to extract the Orpha numbers that are needed for sharing data at EU-level from Orphadata (www.orphadata.org). For that, the Orphanet-classifications were used. It was decided to base the MF on an algorithm in order to minimise the resources needed for long term maintenance and availability. As the Orpha nomenclature and classifications are kept in a structured way in the Orphanet database, the algorithm will be reproducible at any time given that there will be no major restructuring of the Orphanet Database. It will also allow the extraction of the MF in different languages automatically reducing the resources needed at country level.

The classification consists of four different disjointed levels of Orpha numbers: Category, Clinical groups, Disorders and Subtypes. Each level is further represented by typologies (i.e. “disorders” are represented by 6 different types, including “disease” and “malformation

syndrome" amongst others)[1] and have a decreasing complexity. This means that a disorder can be included in a category or a clinical group. It might be subdivided in different subtypes. The different levels are depicted in Figure 1.

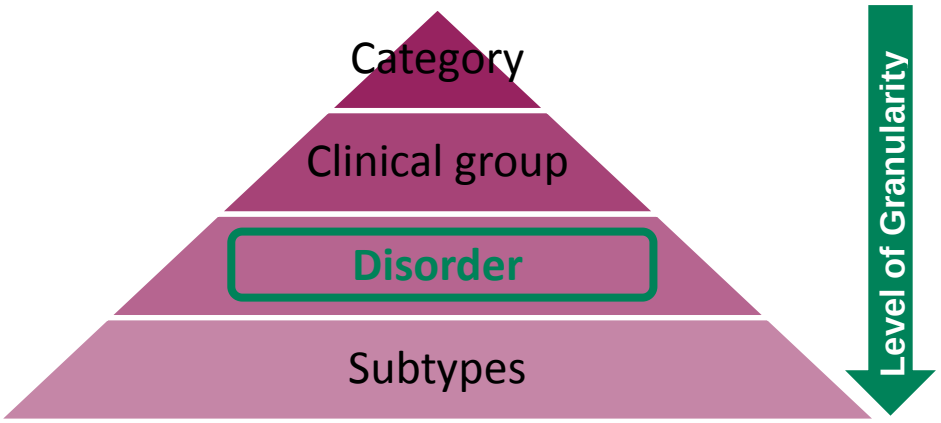

Figure 1: Different levels of the Orphanet classification

The MF was created by only selecting the disorder level. The categories, clinical groups and subtypes were discarded. As well, deprecated[2] and historical[3] entities were discarded.

For the creation of the MF it is also useful to know Orphanet's "ICD-10 coding rules for rare diseases". To understand the structure of the Orphanet nomenclature and why it enables interoperability, it is important to know how the Orphanet nomenclature is produced and how Orphanet attributes ICD-10 codes to Orpha numbers.

A detailed description of the "ICD-10 coding rules for rare diseases" and of the "Procedures: Orphanet inventory of rare diseases" can be found on the Orphanet website (link: see chapter 10).

5. Description of the Master file for statistical reporting with Orphacodes

It is intended that the final version of the MF will be downloadable from the Orphadata website (e.g. from http://www.orphadata.org) in XML-format. For the moment, it is only available on the RD-Action website in Excel-format (www.rd-action.eu). The MF is an extract from the Orphadata and contains only the selected Orpha numbers with the relevant

[1] Disorders include: disease, malformation syndrome, clinical syndrome, morphological anomaly, biological anomaly and particular clinical situation in a disease or syndrome.
Sub-type include: clinical sub-type, etiological sub-type and histopathological sub-type.

[2] "These are entities formerly considered as distinct diseases and now recognised as being part of another disorder present in the database." (Source: http://www.orphadata.org/cgi-bin/docs/userguide.pdf)

[3] "Entries that were described long time ago, most of the time before the genetic era, and for which the princeps article is still available but no further literature exists that confirms their existence" (Source: http://www.orphadata.org/cgi-bin/docs/userguide.pdf)

information that should be used for data sharing at EU-level. For each Orpha number, the following information is available: Orpha number, Name, Synonym and ICD-10-code Version 2016.

The MF in Excel-format is structured as described below. If the XML-file, that is intended to be provided on the Orphadata website, is imported into a spreadsheet software, for example Microsoft Excel, the data will be organized likewise.

1st Spreadsheet (Codes for statistical reporting):

This spreadsheet contains a list of Orpha numbers that should be used for coding rare diseases in Member States. It was generated from the Orphanet classifications using the logic that is described in chapter 4. It only contains the Orpha number and the preferred term of the disease. This spreadsheet is intended for users that have already implemented the full Orphadata in their reporting/coding system and need to compile data for sharing at EU-level. Whenever data was collected in more detail (e.g. by including all subtypes), the data needs to be aggregated for EU sharing. By using the Orphanet-classifications such aggregation can be handled easily.

2nd Spreadsheet (MF):

This spreadsheet contains all Disease names and synonyms in English language for the Orpha number entries that are required for coding rare diseases on the agreed granularity for sharing RD-data on EU level. This list is an extract that was generated from the Orphadata. The structure of the list contains the following columns:

1st Column (OrphaNumber):

This column contains the Orpha number, a consecutively, specific identification number, that has been assigned to a specific disorder. Their specificity should support data sharing at EU-level of RD for statistic purposes. Only a subset of all the Orpha numbers that are available in the Orphanet Nomenclature are contained in the MF.

2nd Column (Name):

This column contains the preferred diagnostic term or name of a given disorder. The Names are adopted from the Orphanet nomenclature. Every name is assigned to one Orpha number (1st column) by Orphanet. (To read more about how Orphanet attributes names to rare disorders, see “Rare disease nomenclature in English”; link in chapter 10 below)

3rd Column (Synonym):

This column contains the terms considered by Orphanet as perfect synonyms of the term listed in the column Name (2nd column) and that are attached to the same Orpha number. Each Orpha number can have none to multiple synonyms. The number of synonyms can differ from one language to another. The language of synonyms depends on the chosen language.

4th Column (ICD-10 V2016):

The column contains the ICD-10-WHO code in its most recent Version (Version 2016) that was mapped by Orphanet to a disorder, a group of disorders or subtypes, according to the Orphanet ICD-10 coding rules (link: see chapter 10). The number of ICD-10 codes can differ from one to multiple codes.

In general, the number of lines for one Orpha number depends on the number of Synonyms that are appendant to it (see table Table 1).

Table 1: Variation of lines for different Orpha numbers depends on the number of Synonyms

Orpha Number	Name	Synonym	ICD-10 V2016
213	Cystinosis	Protein defect of cystin transport	E72.0
213	Cystinosis		E72.0
15	Achondroplasia		Q77.4
95	Friedreich ataxia	FA	G11.1
95	Friedreich ataxia	FRDA	G11.1
95	Friedreich ataxia		G11.1

3rd Spreadsheet (MF - German example):

This spreadsheet shows an example on how integration of a national coding system into the MF can be done. It shows a sample of codes from the ongoing mapping of the MF and the coding system ICD-10-GM, Version 2018 that is currently used in Germany to code morbidity. It shows additional columns for the translated disease name as well as for the national morbidity code. In this case it is the ICD-10-GM code. This spreadsheet will not be maintained at international level but has to be updated and maintained by the country. The translated disease names can be exported from the Orphadata in case the relevant translation is available.

For implementation into national coding systems it is not necessary to use the English disease names and ICD-10-WHO V2016 codes. They are presented in the German example for demonstration purpose only. If a country uses the ICD-10-WHO Version 2016 there is an existing translation in the Orphadata. The file used for integration into a national coding system can be structured just like the MF itself.

6. Using the Master file for statistical reporting with Orphacodes

It is intended that the MF will be downloadable from the Orphadata website in XML-format in the future. The already existing approaches on how to handle and work with the Orphadata (product 1 or Rare diseases and cross-referencing) could be adapted to the MF because of the similar structure of both files. Due to the fact that the MF is extracted from the Orphanet Nomenclature it only contains information provided by Orphanet. Only the translated disease names (if they cannot be extracted from the Orphadata) and national morbidity codes need to be added to integrate the file into routine coding systems. The codes that are included in the MF represent the minimal number of Orpha numbers that should be used for data sharing of RD in MS in order to support the retrieve of statistical data at this level.

Countries should be encouraged to integrate the file within their national coding files and to make sure that if a national coding system is to be used together with the Orpha numbers, both systems are aligned as much as possible to maximise standardisation of results and to help the coders to find the right code from both systems in one coding step. The addition of the different morbidity coding systems from all countries into one MF has proven not to be feasible in the WP5 Session of the Annual Meeting of the RD-Action in 2016. The participants encouraged WP5 to provide an easier way of presenting and working with the MF. Therefore an approach of centrally maintaining all MFs will not be followed further in this WP. It can be considered again, if necessary, in the future, as all prerequisites for this approach are in place.

7. Guidelines

In the framework of Deliverable 5.2 (Standard procedure and guide for the coding with Orphacodes), the WP5 has already published guidelines (link: see chapter 10), which are highly recommended to be used for the routine coding of RD. This procedure will create consistency and a standardized level of quality of coding as well as minimize the coding effort for users and the effort of statistical evaluations.

The guidelines extracted from D 5.2 are the following:

Guideline 1 - Several tools and strategies could be set at MS level to produce data or statistics for RD, nevertheless each country should set this strategy accordingly to a standard principle of maximizing exhaustiveness as well as possible re-use of existing data collections.

Guideline 2 - Code the data in a way that the reporting can compile to the granularity of the international recommended list of Orphacodes (MF-granularity). If no further national needs for reporting are necessary, use the codes from the MFdirectly.

Guideline 3 - Whenever possible capture the information of the diagnostic assertion for all cases. Use the Options “Suspected rare disease”, “Confirmed rare disease” and “Undetermined diagnosis”. Additional options might be helpful.

Guideline 4 – Although rare disease registries (disease, population or patient based) should promote the use of data standards to increase interoperability of their data, they should not be the only instruments upon which the EU strategy to produce health statistics for RD at population level relies.

Guideline 5 - Update your coding resource according to the internationally agreed cycle in order to have the most recent coding file and to ensure comparability.

Guideline 6 - If Orphacodes are used together with another national coding system for morbidity coding, the two systems should be linked in a standardized way to ensure that code combinations are standardized and the coding effort for the user is minimized.

8. Release schedule

The MF should provide a frame for the data sharing at EU-level. This could be achieved through versioning of the file and regulation of the update cycle. With this, also the routine coding process could become more practicable and stable. Regular updates of the coding files are necessary as the field of rare diseases is evolving fast. Still, an agreed update cycle that is followed by all countries is necessary to ensure, that all countries use the same version. From experiences with ICD it is recommended to have an annual cycle of updates. The WP5 team agreed that the MF should be published (updated) once a year. The publication of a file that should be used for data sharing at EU-level in a more frequently rhythm could result in a bureaucratic burden for the implementation and lead to inconsistences in the captured data.

The versioning information should be captured together with the data collected. Every year the list of codes (MF) should be provided for implementation in a standardized way. Ideally this would be aligned to ICD updates (File to be used from January first onwards for one year) as the two coding systems will be used together in the same settings in many EU-countries. This presented file has to be implemented by the countries as well in order to cater to the frequent updates in the Orphanet Nomenclature.

Every year with the release of a new file the old file should be marked as old version but should be kept available online for historical reasons and as reference. Ideally a list of changes should be provided with the new file in order to ease implementation for the users (e.g. adaptation of the national morbidity codes).

9. Content development in the testing phase

As foreseen in RD-Action WP5 work plan for the third year of the RD-Action, the current beta version of the MF will be field-tested and feedback will be taken into account for the publication of a final version at the end of the RD-Action. The final version will be further published annually by Orphanet. This guideline document will be also updated according to the feedback from the field-testing phase.

If you wish to propose the addition of a specific Orpha number to the MF for the final version of the MF please send an E-Mail including your proposal to the WP5 team. It will review the proposal together with the Orphanet team and decide if the Orpha number will be added to the list of codes for data sharing at EU-level. If so, it will be discussed how the algorithm for generating the MF can be adapted to include this Orpha number. If possible the Orpha number will then be included in the next version of the MF.

If you wish to contact the WP 5 team, please contact rare-diseases@dimdi.de.

10. Further information on the topic

To learn more about the work of the RD-Action please visit the RD-Action website (www.rd-action.eu).

To learn more on how the Orphanet nomenclature is produced, see “Procedure: Orphanet inventory of rare diseases” (http://www.orpha.net/orphacom/cahiers/docs/GB/eproc_disease_inventory_PR_R1_Nom_04.pdf).

To read more about how Orphanet attributes names to rare disorders, see “Rare disease nomenclature in English” (http://www.orpha.net/orphacom/cahiers/docs/GB/eproc_Disease_naming_rules_in_English_PR_R1_Nom_01.pdf).

To learn more about the Orphanet ICD-10 coding rules for rare diseases please visit the Orphanet website (http://www.orpha.net/consor/cgi-bin/Education.php?lng=EN).

To learn more about the Orphadata and its content, see http://www.orphadata.org/cgi-bin/docs/userguide.pdf.

To learn more about the work of the WP5, especially the Deliverables (D) and Milestones (MS) please choose from the links below:

- RD-Action WP5 State of the Art survey
- MS24: Specifications of an integrated coding application with Orphacodes
- D5.1:Review document of existing technical implementations for RD coding of MS
- D5.2: http://www.rd-action.eu/wp-content/uploads/2017/05/D5.2_Standard-procedure-and-guide_final.pdf