“Final report - Spain”

[31/12/2021]

This report is part of the project 826607 RD-CODE which has received funding from the European Union’s Health Programme (2014-2020).

The content of the report represents the views of the authors only and is their sole responsibility; it cannot be considered to reflect the views of the European Commission and/or the Consumers, Health, Agriculture and Food Executive Agency or any other body of the European Union. The European Commission and the Agency do not accept any responsibility for use that may be made of the information it contains.

TABLE OF CONTENTS

1 – Tasks and timeline

The Center for Biomedical Network Research on Rare Diseases (CIBERER) and the Foundation for the Promotion of Health and Biomedical Research in the Valencian Region (FISABIO) took part as participant institutions at the RD-CODE project along with six collaborating stakeholders: Consejería de Salud, Región de Murcia; Servei Català de la Salut (CatSalut); Departamento de Salud del Gobierno de Navarra; Dirección General de Salud Pública de la Junta de Castilla y León; Departamento de Salud, Gobierno Vasco; Dirección General de Salud Pública, Generalitat Valenciana.

The overall goal was to promote the use of the Orphanet nomenclature for implementation into routine coding systems by establishing equivalences between ICD-10-ES and Orphacodes. This would enable a standardized and consistent level of information to be shared at Spanish and European level. In particular, for Spain, the goal was to pilot the implementation of Orphacodes according to the "Standard procedure and guide for the coding with Orphacodes" and the "Specification and implementation manual of the Master file" at the rare diseases (RD) registries in Spain.

The Spanish participation in the WP4 of the project was originally divided in 4 phases: the pilot implementation phase from m1 to m12; the preliminary results analysis from m13 to m15; the extension of the use of Orphacodes from m15 to m28; and the final report, including recommendations for the implementation of Orphacodes, from m28 to m30. This schedule was later modified by an amendment of the grant agreement including a six month extension of the project. Outcomes from phases 3 and 4 were shifted to m34 and m36 respectively (Figure 1). The pilot phase, involving the 6 RD regional registries of the collaborating stakeholders' institutions, included the mapping of Orphacodes to the Spanish version of the ICD-10-CM (ICD-10-ES) and the implementation of the Master file for statistical reporting with Orphacodes (MF) using as template the file available from the RD-ACTION website. The second phase involved the analysis of the results obtained during the pilot phase and the report compiling the advances achieved. A third phase which was aimed to extend the regional network of RD registries was carried out in order to cover a higher ratio of the Spanish population to be coded with Orphacodes. The fourth and final phase was meant to collect and discuss the results of the previous phases, as well as to elaborate recommendations for the implementation of Orphacodes at the national level, in order to summarize the project achievements in Spain.

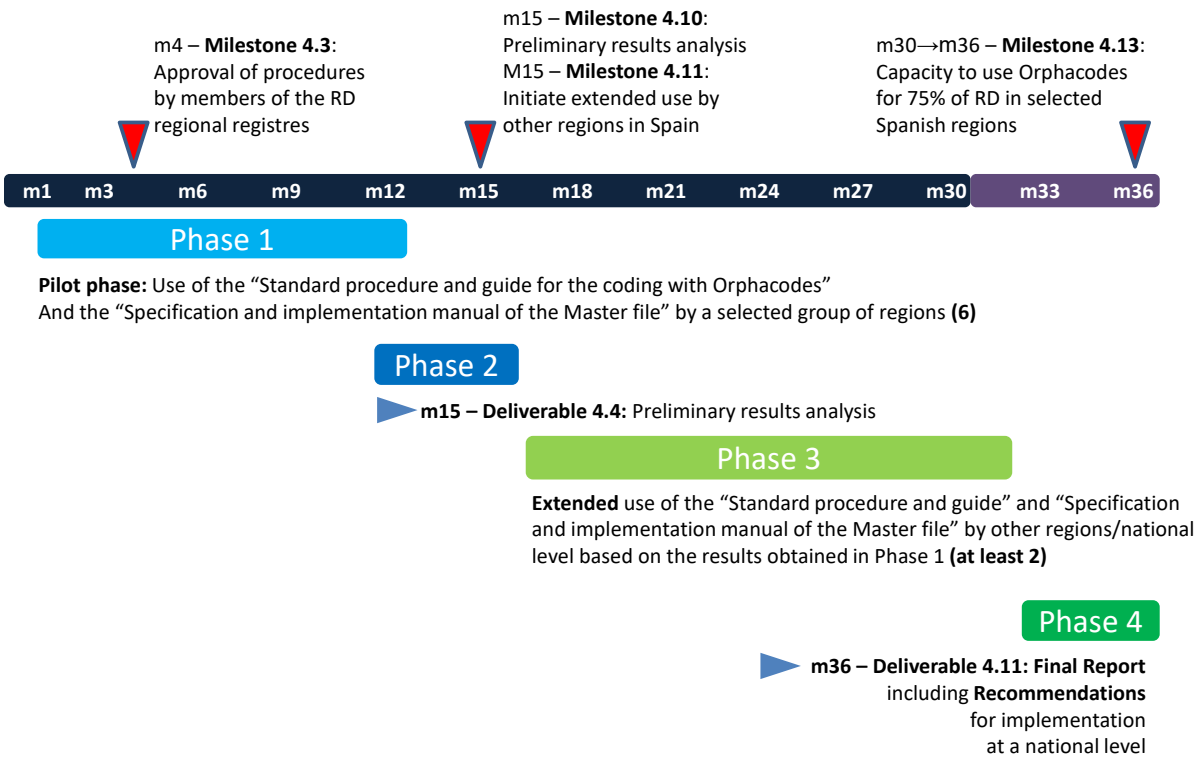

Figure 1. Extended timeline for the Spanish participation at the WP4 of the RD-CODE project.

2 – Collaborating stakeholders, tools and strategies

The RD regional registries from the Basque Country, Castile and Leon, Catalonia, Murcia, Navarre and Valencia Region covering around 40% of the Spanish population were involved from the beginning of the project and participated in the pilot implementation phase. The results from the implementation phase were summarized and discussed in the deliverable 4.4: Preliminary results analysis and "Adaptation of procedure for Orphacodes use in Spain" which can be downloaded from the Health Programme DataBase - European Commission (Item D.10).

Following the pilot implementation phase and the report of the preliminary results obtained during such, the extension phase was initiated to widen the coverage of the Spanish population by other RD regional registries joining the project. The representatives from several regional registries, including Galicia, La Rioja, Baleares, Extremadura and Madrid were contacted in March/April 2020 to inquire about the potential participation of their RD registries in the third phase of the RD-CODE project and to establish some common guidelines, goals and chronogram in order to achieve implementation of Orphacodes in these regions. However, although most of them expressed their interest in joining the project, the impact of the COVID-19 pandemic over the resources that these registries could dedicate to accomplish the set goals, prevented them from participating, with the exception of the Madrid RD registry, thus reaching a final coverage of 54% of the Spanish population.

The «Standard procedure and guide for the coding with Orphacodes» and the «Specification and implementation manual of the Master file» both developed in the frame of the previous Joint Action on Rare Diseases RD-ACTION (2015-2018) were used as reference and the «Master file for statistical reporting with Orphacodes» also produced in the frame of the RD-ACTION and updated with the ICD-10-ES equivalences within the RD-CODE project was the vehicle to homogenize Orphacodes assignment throughout the Spanish territory.

3 – Extended regional network and collaborators

As mentioned above, Madrid joined the six regions originally involved in the project: Castile and Leon, Catalonia, Murcia, Navarre, Basque Country and Valencia Region. The seven regions had a previously established RD regional registry with intrinsic features. Specifications of the regional registries characteristics corresponding to the six registries originally involved in the project can be found in deliverable 4.4 Health Programme DataBase - European Commission (Item D.10). The details of the Madrid registry are as follows:

General information

The Madrid registry (SIERMA) was created in 2015 and is managed and maintained by both the registry's team and an external enterprise. The computing system is based on Microsoft's SQL Server as Database Management System (DBMS). The system uses the SQL Server Relational Database engine for data storage and management, the SQL Server Reporting Services for design reports, the SQL Server Analysis Services (Multidimensional / Tabular) to design an analytical system and the SQL Server Integration Services for extracting data. At the beginning of the project, the registry included, 398,836 individuals with at least 1 RD (candidate patients from ICD codes – before validation), and 19,298 cases validated and confirmed for the 2010-2018 period.

Information sources

The sources of information of this registry are: the hospital discharges database (CMBD), the primary care electronic health record, the orphan drugs registry, the mortality registry, the Madrid renal diseases registry (REMER), the newborn screening for endocrine and metabolic diseases, the RD Research Centre ISCIII patients' registry, the clinical registries of selected diseases, and direct notification by specialists through an integrated module at the hospital electronic health records.

Coding systems in use

The input coding systems in use by the SIERMA are: ICD-9-CM, ICD-10, ICD-10-ES (from 2016) and ERA-EDTA. Cases coded with SNOMED-CT and Orphacodes are also received from the ISCIII patients' registry.

Moreover, a few additional collaborations were also evaluated to take advantage of the momentum generated by the RD-CODE project and the implementation of Orphacodes in Spain. At the hospital level, the Sant Joan de Déu (HSJD) hospital from Barcelona got in

contact to study the synergies between their registry and the potential exploitation of the MF developed during the project. This collaboration has provided a different perspective to the use and implementation of Orphacodes due to the nature of the registry (hospital level) and the purpose that it aims to fulfil. Details regarding to the collaboration with the HSJD registry can be found in section 5.

As far as the issue of undiagnosed patients, some meetings were conducted with representatives from the undiagnosed patients programme ENoD (CIBERER) in order to pilot different coding options for those cases that are suspected to suffer from a RD but for which a diagnosis hasn't been reached. A first attempt with a selected subset of cases coded with the HPO classification system was performed with the aim of identifying them within RD registries. However the dismissal of a pilot trial by WP5 leaders due to expected limitations of the study, and the final decision to create a unique dedicated Orphacode to identify undiagnosed RD patients put this collaboration on hold.

4 – Update and maintenance of ICD-10-ES correspondences for Orphacodes

The MF resulting from the RD-ACTION joint action was originally extracted from the 2018 version of the Orphanet nomenclature. New versions of the Orphanet nomenclature are yearly released thus changing the core of the aggregation level (disorder level entities). In a similar way, the ICD-10-ES is also revised and modified every two years. These regular updates of the codification systems employed to identify RDs make it essential to perform parallel updates of the MF and the cross-references that it provides in order to preserve its validity to be used at Spanish RD registries (Zurriaga *et al.* 2021) – see publications list at section 6.

After the pilot implementation phase, the coverage of the original MF Orphacode/ICD-10-ES cross-referencing increased to 5,664 (Cavero-Carbonell *et al.*, 2020; Rico *et al.*, 2020; Rico *et al*., 2021) out of the 5,775 listed in the MF. The 111 Orphacodes lacking ICD-10-ES equivalence were double-checked at the API - RD-CODE and the MSCBS RD visor, and submitted through the Helpdesk - RD-CODE tool. Feedback from BfArM colleagues in Germany was received and the three sources were combined to solve 22 additional equivalences.

At the beginning of 2021, the MF extracted from the Orphanet nomenclature released in 2020 was compared with the one from 2018 and it was found that 168 Orphacodes had been removed while 551 new codes had been added. An additional 125 codes had been also removed from the ICD-10-ES and 739 and 442 had been added and modified, respectively. All equivalences were updated to the ICD-10-ES 2020 version and 264 of the new Orphacodes from the 2020 MF were included with their corresponding ICD-10-ES equivalences, reaching a total of 5,787 (94% coverage).

Later on, the 2021 version of the Orphanet nomenclature, along with its corresponding MF containing 6,177 Orphacodes, added new changes that had to be taken into account. Yet

again, 39 Orphacodes were either inactive or classified as group or subtype. Furthermore, 63 new Orphacodes were incorporated and 489 had a new or different ICD-10 cross-referencing. The ICD-10-ES mapping of all the Orphacodes affected by these modifications was revised and 433 were validated, modified or incorporated. A total of 5,873 disorder level Orphacodes from the 2021 version of the Orphanet nomenclature were mapped to ICD-10-ES reaching coverage of around 95% of all disorders listed in the current version of the MF (Figure 2).

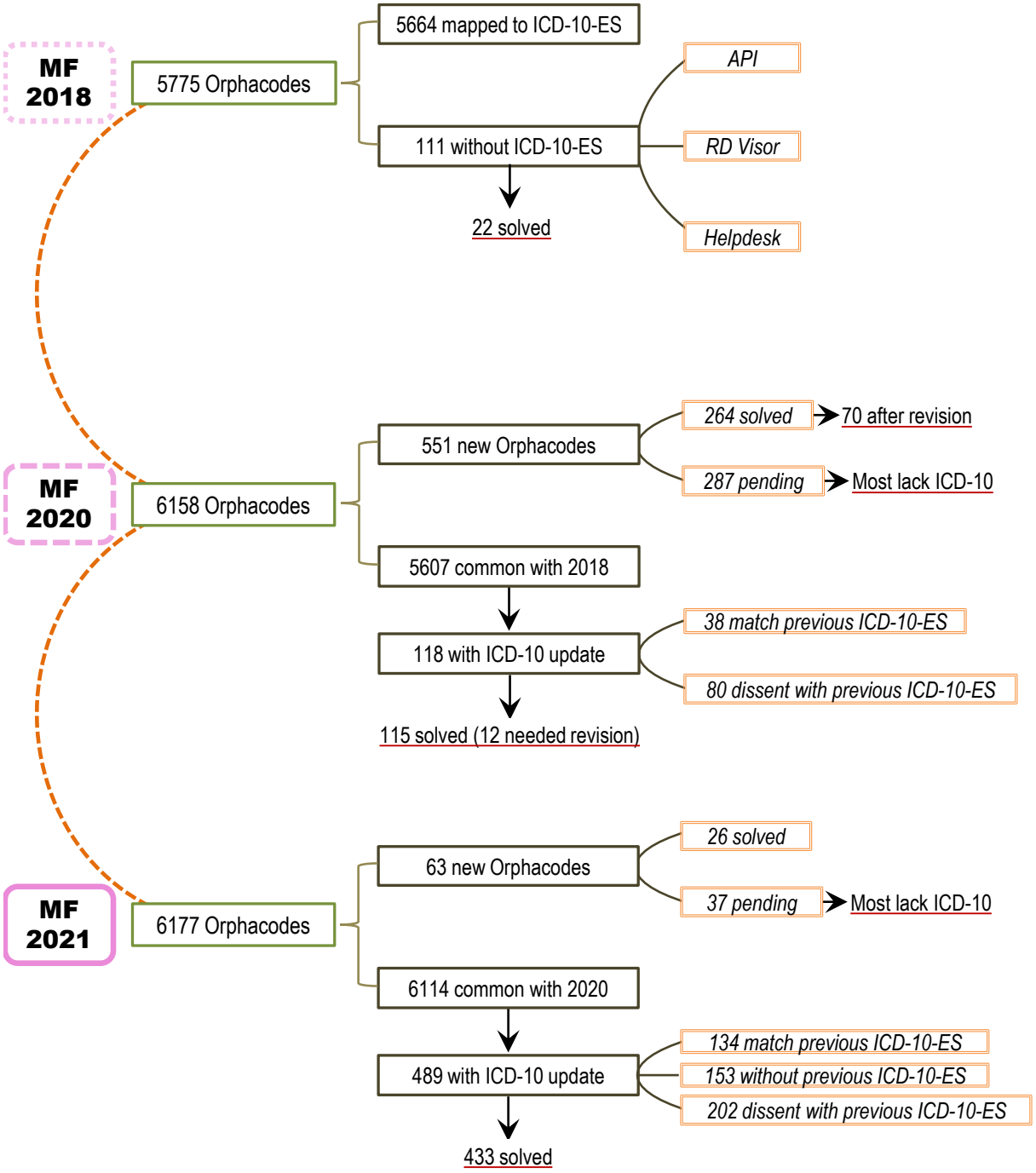

Figure 2. Schematic representation of the update and evolution of the MF and the ICD-10-ES cross-referencing progress during the RD-CODE project.

Complementary approaches

In addition to the ICD-10-ES classification, there exists an ICD-based coding system developed to better identify congenital anomalies (ICD-10-BPA). This coding system is in use at the congenital anomalies registries of several regions of Spain and its use is a requisite to be full member of the European network of population-based registries for the epidemiological surveillance of congenital anomalies (EUROCAT). The ICD-10-BPA classification is, in essence, an extension of the Q chapter of the ICD-10-CM (and its Spanish version ICD-10-ES).

As the results of the mapping to ICD-10-ES showed an aggregation of Orphacodes at the Q chapter of the ICD-10-ES often leading to highly redundant equivalences (Cavero-Carbonell *et al.*, 2020; Rico *et al.*, 2020; Rico *et al.*, 2021), an attempt to alleviate this low specificity was made by establishing cross-references to ICD-10-BPA. To make an assessment of the potential improvement that the mapping of Orphacodes to ICD-10-BPA could provide, we started by mapping 1,047 Orphacodes that were mapped to the most redundant ICD-10-ES codes from the Q chapter: Q04.3, Q82.8, Q87.0, Q87.1, Q87.2, Q87.89, Q92.2 and Q93.5. The ICD-10-BPA allowed us to establish some additional one-to-one cross-references (15) even if the vast majority of the Orphacodes remained clustered within highly unspecific ICD codes. The benefits also varied depending on the type of disorder and the original ICD-10-ES. Regarding the Orphacodes originally mapped to the Q04.3 code of the ICD-10-ES, 5 ICD-10-BPA codes were available (Figure 3). On contrary, for the codes mapped to the Q93.5 code of the ICD-10-ES, no additional code was found at the ICD-10-BPA. Overall, the disorders with the best resolution improvement by the ICD-10-BPA mapping were those originally assigned to the Q87.0, Q87.1 and Q87.2 codes of the ICD-10-ES with 23 alternative ICD-10-BPA codes mapped to 58 of the 259 (>22%) Orphacodes previously mapped to these 3 ICD-10-ES codes (Figure 3).

Regarding absolute numbers, 2,115 cross-references were established between 538 ICD-10-BPA codes (≈50% of the ICD-10-BPA classification) and 1,843 Orphacodes. In conclusion, the ICD-10-BPA classification is yet lacking specificity for a large number of congenital RDs. However, cross-referencing between ICD-10-BPA and Orphacodes seems sensible not only because the congenital anomalies regional registries are an important source of information for the Spanish regional RD registries, and ultimately for the national registry, but given that, as shown above, this classification facilitates the identification of at least 50 disorders which are virtually impossible to discern using the ICD-10-ES.

In order to optimize the output of the cross-referencing between ICD-10-BPA and Orphacodes, the final set of ICD-10-BPA codes provided to the SIER-CV were those that, by consensus, are used to code syndromic congenital anomalies by the EUROCAT consortium. The congenital anomalies registry of the Valencia Region is a full member of EUROCAT and thus adopts its criteria for the codification of congenital anomalies cases. The congenital anomalies registered with these ICD-10-BPA codes were mapped to 118 Orphacodes that will, in the end, be used to generate the corresponding RD cases at the RD registry's software.

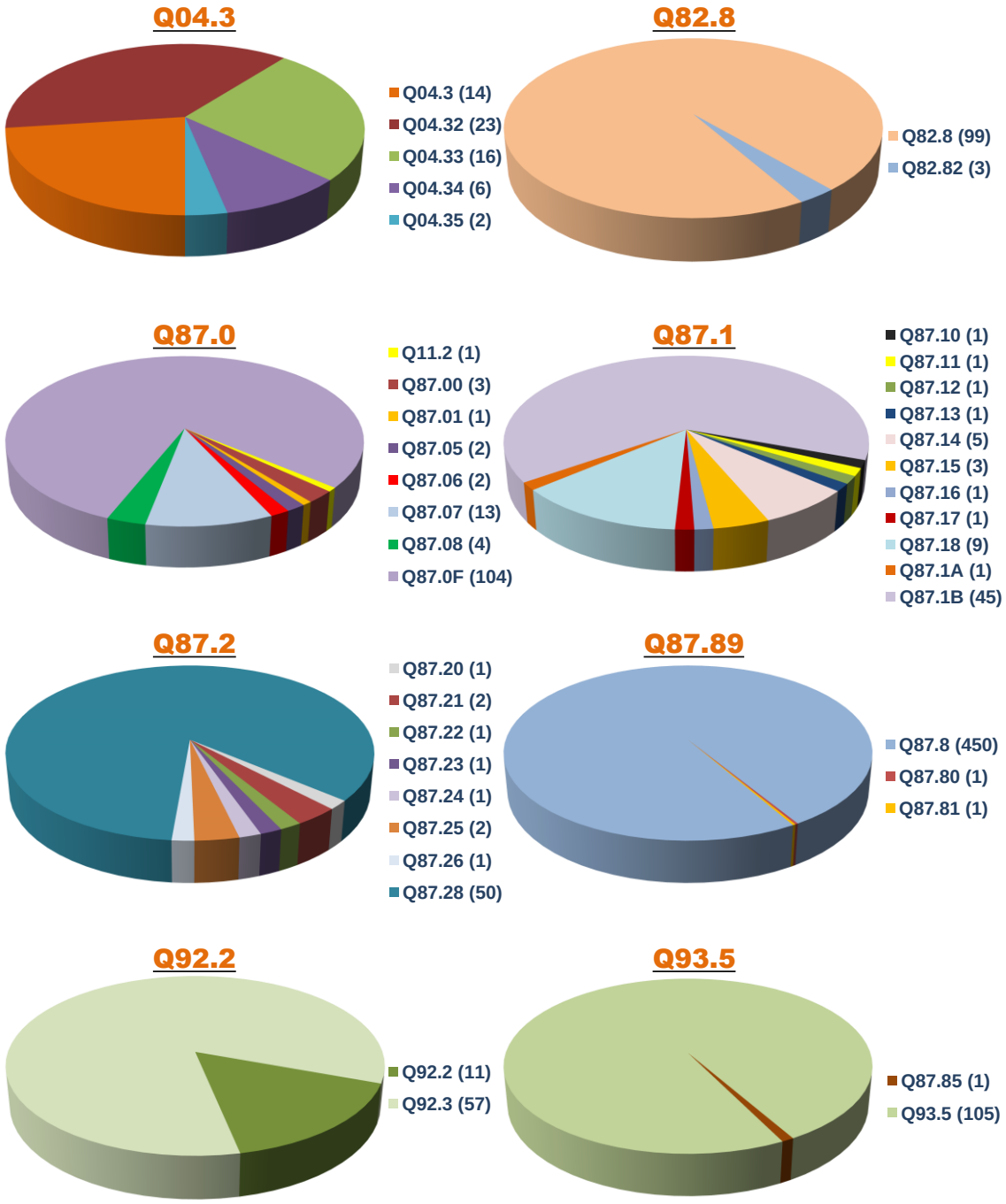

Figure 3. Comparison of the specificities of the ICD-10-ES (dark orange - underlined) and the ICD-10-BPA (dark blue) for selected congenital disorders.

5 – Feedback and advances from partners

Madrid

Goals:

- Update the ICD-10/ICD-10-ES SIERMA catalogue of RDs using the MF.
- Implement MF mapping among Orphacodes and ICD-10/ICD-10-ES at electronic clinical health records revisions.
- Incorporate or partially update a subset of the MF to their RD catalogue of direct notification by specialists.
- Provide experience-based feedback in complex cross-references.

Results:

- 13 codes incorporated to the ICD-10/ICD-10-ES SIERMA catalogue of RDs (Table 1).
  - 487 additional ICD-10-ES codes under evaluation.
- 42 Orphacodes used for validation at electronic health records revisions (Table 2).
- 13 RDs were incorporated to their RD catalogue of direct notification by specialists (Table 1).
  - A total of 4,828 disorders or groups of disorders are anticipated to follow soon.
- Takotsubo Syndrome ICD-10-ES code was modified in the MF after feedback received from SIERMA.
  - Suggestions for 119 additional cross-references were also provided.
- 403 different Orphacodes used to identify confirmed RD cases since 2015.

Table 1. New and updated rare diseases at ICD-10/ICD-10-ES SIERMA catalogue.

Rare disease or group of diseases	**ICD-10-ES**
Other neuromuscular dysfunction of bladder	N31.8
Other disorders resulting from impaired renal tubular function: distal renal tubular acidosis	N25.89
Secondary polycythemia	D75.1
Acquired coagulation factor deficiency	D68.4
Other specified deforming dorsopathies: Camptocormia	M43.8
Transient hypogammaglobulinemia of infancy	D80.7
Other thalassemias: (Hb-C) thalassemia, Hb Lepore-beta-thalassemia syndrome	D56.8
Other specified disorders of brain: Symmetrical thalamic calcifications, Infantile choroidocerebral calcification syndrome, ANK3-related intellectual disability-sleep disturbance syndrome	G93.89
Lambert-Eaton syndrome in neoplastic disease	G73.1
Isolated myocarditis: Idiopathic giant cell myocarditis	I40.1
Other cardiomyopathies: Familial isolated arrhythmogenic right ventricular dysplasia, Cirrhotic cardiomyopathy, Left ventricular noncompaction	I42.8
Respiratory bronchiolitis-interstitial lung disease syndrome	J84.115
Epileptic spasms: West Syndrome	G40.82

Table 2. Rare diseases validated with help of the RD-CODE MF.

Rare disease	Orphacode
Alport syndrome	63
Cockayne syndrome	191
Cornelia de Lange syndrome	199
Congenital factor XI deficiency	329
Simpson-Golabi-Behmel syndrome	373
Gorlin syndrome	377
Oculocerebrorenal syndrome of Lowe	534
Monosomy 21	574
Noonan syndrome	648
Isolated Pierre Robin syndrome	718
Rett syndrome	778
Rubinstein-Taybi syndrome	783
Seckel syndrome	808
Silver-Russell syndrome	813
Smith-Lemli-Opitz syndrome	818
Smith-Magenis syndrome	819
Sotos syndrome	821
Stickler syndrome	828
Turner syndrome	881
Tetrasomy 12p	884
Von Willebrand disease	903
Williams syndrome	904
Aarskog-Scott syndrome	915
Coffin-Siris syndrome	1465
Monosomy 18p	1598
1p36 deletion syndrome	1606
2q24 microdeletion syndrome	1617
Jacobsen syndrome	2308
3M syndrome	2616
Costello syndrome	3071
Weaver syndrome	3447
Infantile spasms syndrome	3451
Monosomy 22q13.3	48652
Congenital muscular dystrophy, Ullrich type	75840
Distal monosomy 4q	96145
8q22.1 microdeletion syndrome	178303
1q21.1 microdeletion syndrome	250989
8p23.1 microdeletion syndrome	251071
Proximal 16p11.2 microdeletion syndrome	261197
Arnold-Chiari malformation type I	268882
Primary Sjögren syndrome	289390
Developmental delay-facial dysmorphism syndrome due to MED13L deficiency	369891

Castile and Leon

The updated status of the implementation of Orphacodes at RERCyL -Rare Diseases Registry of Castile and Leon- can be summarized as follows:

5,772 Orphacodes are included in their system, from which 1,047 Orphacodes include the ICD-10-ES correspondence in their software, and 539 have been used to identify confirmed RD cases.

The MF resulting from the RD-CODE project is uploaded at the software and is available to make consultations, research and analysis, and constitutes also the main source of information to establish the equivalences finally included in their system. This way, the upload of the information contained at the MF is made according to the registry requirements and in a progressive and continuous manner.

This registry is also working towards the implementation of SNOMED-CT to ICD-10 equivalences in order to fulfil the requirements of the Spanish Ministry of Health.

Most recent updates:

- Possibility of browsing diagnoses by Orphacode and displaying this codification in search windows and general information of the patient at the RERCyL's software.
- Assignment of Orphacodes to cases retrieved from ICD-10 based information sources using the MF developed at the RD-CODE project, taking into consideration the validation guidelines provided by the national registry of RDs as well.

The aim for the last year of the project was to keep track of the changes and update both the Orphacode list and the correspondences to ICD-10-ES taking into account the MF updates from the RD-CODE, keeping in mind the needs and capabilities of the registry.

In addition, the materials generated during the RD-CODE project have been shared with the validation unit and Castile and Leon RD diagnosis unit in order to keep homogeneity in the communication of cases to the registry. This unit is the only source that communicates the cases already with an Orphacode. When the selection of an Orphacode is not straightforward, full diagnosis description is submitted and the code is assigned at the registry.

Overall, the inception of an Orphacode codification culture has been set. This codification has unquestionably helped with the management of RD cases, and, in many of them, has provided further specificity. Whenever the ICD-10-ES is more specific, the Orphacodes have proven to be useful as a vehicle to communicate with other registries, including the national RD registry, in a standardized manner.

Currently, when validating new ICD-10 based cases:

- If there is a 1:1 correspondence to an Orphacode in the MF, this Orphacode is added to the diagnosis.
- In case there are several Orphacodes corresponding to the ICD-10-ES input code, the most likely Orphacode is displayed along with the diagnosis but a 'click and select' list

containing all the Orphacodes matching this ICD-10-ES is also provided in order to change it after validation of the case, if necessary.

- This is still a beta version for a reduced number of diseases.
- Example:

ICD-10-ES code M31.0 to which the Goodpasture syndrome is associated. This ICD-10-ES code is quite generic and is commonly correlated with the Orphacode 889 (Cutaneous small vessel vasculitis). When the M31.0 code is received by the registry, the Orphacode 889 is initially assigned but the Orphacode 375 is available at the drop-down menu in order to be selected if the diagnosis of Goodpasture syndrome is confirmed.

The following actions will be pursued in the mid-term:

- Concluding the implementation of the ICD-10-ES/Orphacode correspondences paying close attention to those with multiple equivalences.
- Translate terms into Spanish to avoid misunderstandings.
- Continue promoting the use of Orphacodes in the registry.

Murcia Region

- The full MF has been uploaded to the registry's system.
- Currently, 16,288 validated RD cases are associated with at least one Orphacode.
- A total of 1,035 different Orphacodes have been used to register the corresponding validated RD cases.
  - Not all the Orphacodes were chosen from the MF.
    - Orphacodes corresponding to a "subtype of disorder" level were used in those cases for which the information retrieved was enough to reach this level of specificity.
      - Validated cases do not lose specificity and can always be aggregated for statistical analysis.

Basque Country

- Cases are retrieved by direct notification from the clinicians using a drop-down list of RD diagnoses.
- 974 different Orphacodes (belonging to the "group of disorders" and "subtype" levels) have been used to identify confirmed RD cases.

Catalonia

- The cases are retrieved by direct notification from clinical experts on RDs from several medical specialties.
- 1,253 different Orphacodes have been registered to identify at least one RD confirmed case.

Navarre

- Catalogue of ICD-10/ICD-10-ES codes in use to identify RD cases updated according to the MF.

- 284 different Orphacodes were used to classify 1,125 RD cases retrieved from the Congenital Anomalies and Hereditary Diseases Registry of Navarre (RACEHNA).
- At least one episode was received from the CMBD, during the period 2016-2020, for 507 out of the 1,125 RD cases identified from RACEHNA.
    - The diagnoses of 329 of these RD cases, 140 different Orphacodes, were identified with an ICD-10-ES code listed at the MF.
        - The ICD-10-ES codes with which the RD cases were registered at the CMBD matched the ICD-10-ES of the MF for 71 Orphacodes.
- This registry has, so far, identified 937 RD cases that fall within the current scope of the Spanish national registry of RDs (REER), and that are also coded with Orphacodes.
- Since the beginning of the RD-CODE project, every RD case under study is assigned an Orphacode (at the disorder or subtype level) after validation through the revision of its clinical health record.
- The implementation of the MF resulting from the RD-CODE project at the hospital clinical health records software was proposed:
    - A pilot at the Neurology service of the Navarre´s University Hospital was being planned but the pandemic put it on hold. It will hopefully be resumed in the near future.

Valencia Region

During the RD-CODE project, a number of strategies to efficiently implement Orphacodes at the SIER-CV through the MF have been discussed and tested. Following this optimization process the most relevant conclusions and results reached are:

- The MF comprises RDs that do not match the SIER-CV case definition (e.g. Neoplasia, infectious diseases, etc.). This circumstance itself is enough to hinder the automation of the implementation of Orphacodes since the establishment of an adequate filtering system needs to be rendered. Close to 30% of the Orphacodes listed at the MF have been included at the SIER-CV software.
- In order to allow the configuration of the diagnoses at SIER-CV with a unique combination of codes, a codification triad combining ICD-10-ES, Orphacode and SNOMED-CT was established as a requirement for case construction. This approach also helped to solve the cases in which a narrower Orphanet term was mapped to a broader ICD-10-ES code. However, as the MF does not provide the mapping to SNOMED-CT, this had to be individually and manually selected by the SIER-CV personnel. As many as 1,734 Orphacodes have been implemented within these codification triads, and 539 (including 209 automatically assigned) of them have been already used to register 16,183 confirmed RD cases.
- The fact that the MF is only distributed in English has not been an inconvenience as the implementation has been carried out manually due to the limitations described above. Nevertheless, if the automation of the Orphacodes implementation is at some point feasible, it would require the Spanish version of the MF as this is the language used by the software.
- Despite the limitations encountered, the MF has allowed the SIER-CV to double the number of RDs identified in its system when compared to the status previous to the

project. The MF has significantly helped in the generation of the ICD-10-ES + Orphacode + SNOMED-CT codification triads. A master table containing the combination of these three classification systems for each diagnosis has been created and allows the right identification of the cases that are subsequently submitted to the national registry.

- The cross-references established in the MF have enabled the retrieval of new and more specific ICD-10-ES codes from the information sources and, in those cases in which an exact mapping has been established, a system to automatically validate the cases will be possible in the future, increasing both the number and the quality of the RD cases registered.

All these efforts, advances and outputs can be summarized as follows:

- The RD-CODE project, through the MF, has improved the retrieval of cases from CMBD by looking for specific ICD-10-ES codes.
- The validation of RD cases is more efficient now that at the beginning of the project.
- The MF implementation has allowed to reach higher level of granularity in case validation and coding.
- The implementation of the MF is expected to unlock the development of a notification process from the Reference Centers, Services and Units of the National Health System (CSUR-SNS) as well as from specialized and coordination practices and general assistance for RDs.
  - This would allow the use of Orphacodes at the information source, avoiding the information loss in those cases in which ICD-10-ES is more unspecific.

<u>Sant Joan de Déu</u>

Goals:

- Improving the identification of patients diagnosed of RDs in their electronic health records system (HCIS).

- Implementing Orphacodes, along with their ICD-10-ES cross-references, to code RD diagnoses.

- Identify undiagnosed RD patients (SWAN).

Methods and strategies:

- Creation of a pediatrics RD diagnoses catalogue based on the Orphanet nomenclature and its inclusion in their HCIS.

  - Initial distribution of disorders by clinical service based on the Orphanet classification.

  - Implementation of the information available to filter the entries registered (e.g. Classification level).

- Selection of pediatric entities that will be finally included at the registry's catalogue (a prerequisite for inclusion at the catalogue is to have at least one case previously diagnosed with that disorder).
- Look for candidate Orphacodes that could be used for SWAN diagnoses.
- Adaptation of the Orphanet nomenclature:
  - In some cases, internal codes that are equivalent to the same Orphacode have been generated at the request of clinicians.
  - Clinical terms have also been complemented with additional information when required.
- Create or adapt an 'ad hoc' cross-references table between Orphacodes and ICD-10-ES (based on the information provided by Orphanet and the RD-CODE project, with the modifications proposed by their clinical documentation service).
- Translation of clinical terms into Catalan.
- Refinement of the entries generated to avoid duplication, inconsistencies and errors in general to generate the ultimate version of the catalogue.

- Establishment of an interdisciplinary working group composed by the heads of clinical areas and professionals from non-clinical departments (information systems, clinical documentation, certifications unit, and informatics).

Current status:

- Pilot of the implementation of the RD catalogue and its features by 4 different hospital services started in October 2021.

- Full development of the RD catalogue completed in November 2021.

- The best possible strategy to update the information available in the catalogue (e.g. incorporate new entries) is yet to be defined.

Major drawbacks:

- Cases for which Orphanet entities lack enough specificity to fulfil the needs of the registry.
  - Internal codes adding specificity are generated:
    - Orphacode 183675: Recurrent infections associated with rare immunoglobulin isotypes deficiency – Internal code 1 (kappa light chain).
    - Orphacode 183675: Recurrent infections associated with rare immunoglobulin isotypes deficiency – internal code 2 (lambda light chain).

- Cases for which a single ICD-10-ES is not enough.
  - Additional ICD-10-ES codes are applied:

- Orphacode 70596: Congenital Epstein-Barr virus infection
  - ICD-10-ES: P35.8 – B27.00

- Solid tumours:
  - Neither Orphanet nomenclature nor ICD-10-ES have been found enough to fulfil the criteria defined by the registry.
    - Alternative classifications are being explored.
      - These classifications will be cross-referenced to the Orphanet nomenclature.
      - The creation of a dedicated working group on this topic is being considered.

- Codification efforts are time and resource-consuming even with the input received (MF) from the RD-CODE project.
  - Additional resources are being claimed.

- The integration of the catalogue within the HCIS took a long and tedious process.

Overall feedback

In order to estimate the number of RDs that can be identified with their Orphacodes at Spanish registries, the list of Orphacodes in use by 6 different regional RD registries (Madrid; Valencia Region; Basque Country; Castile and Leon; Catalonia and Murcia) was first retrieved in June 2021. A list containing a total of 3,201 different Orphacodes was generated after comparing the lists from each region. Over 70% (2,259) of these Orphacodes have been cross-referenced to ICD-10-ES in the frame of the RD-CODE project. However, when this list was further analysed, only 54 of these Orphacodes were found to be shared by all six regional registries and 1,820 are in use only in one of them. This is mainly due to the fact that, to date, only 22 RDs are of mandatory communication to the RD national registry and each regional registry has established its own strategy towards full implementation of Orphacoding taking into account their needs and capabilities. Despite the current limitation to perform national and transnational statistical analysis for a broad number of RDs, the MF distributed to the regional registries has shown its potential to increase the number of Orphacodes that can be implemented at Spanish registries. The readiness provided by the MF and the momentum generated by the RD-CODE project makes it just a matter of time that, with the implementation of additional Orphacodes at each registry, there is an increasing overlap among the lists of Orphacodes in use at the RD regional registries.

As of October 2021, the list of Orphacodes ready to be used at these registries has experienced a slight increase to reach 3,271 codes from which 2,464 have indeed been used to identify at least one confirmed RD case.

6 – Outcomes

Tasks, milestones and deliverables:

Task 4.4: implementation in Spain (Start date: M1 - End date: M36)

Leader: CIBER

Participant: FISABIO

Subtask 4.4.1 Phase 1: Organisation and set up of the pilot phase. (M1-M12).

Use of the "Standard procedure and guide for the coding with Orphacodes" and the "Specification and implementation manual of the Master file" by a selected group of six regions in Spain.

Milestone 4.3 (M4) (ES) Approval of procedures by members of the RD regional registries.

Procedures approved in April 2019

Subtask 4.4.2 Phase 2: Preliminary results analysis (M13-M15)

Following phase 1 of the Project (pilot phase M1-M12), which includes participation of six regions, results will be evaluated in order to make the necessary adjustments, if needed, before approaching other regional registries to participate in the Project.

Mls 4.10 (M15) (ES) Preliminary results analysis.
Deliverable 4.4 (M15) Preliminary results analysis and "Adaptation of procedure for Orphacodes use in Spain" Report (ES): The first six regions progress and results will be analysed and evaluated.

Accomplished and submitted in March 2020

Milestone 4.11 (M15) (ES) Initiate extended use of the "Standard procedure and guide" and "Specification and implementation manual of the Master file" by other regions in Spain.

First contact established in March 2020

Subtask 4.4.3 Phase 3. Extension to the other regions (M16-M34)

The use of the "Standard procedure and guide" and "Specification and implementation manual of the Master file" will be expanded to other committed regions in Spain not participating in the pilot phase based on the results obtained in Phase 1 after having approached them. A procedure will be defined for that purpose.

Madrid regional registry and Sant Joan de Déu hospital registry engaged

Subtask 4.4.4. Phase 4: Final report (M34-M36)

A final Report including an operational manual and Recommendations for implementation at a national level will be issued from the experience gained over the project and will be made publicly available and disseminated to the Health authorities.

Deliverable 4.11 (M36) Final report (ES): this report will include Recommendations for implementation at a national level.

Accomplished and submitted December 2021

Milestone 4.13 (M36) (ES) Capacity to use Orphacodes for 75% of rare diseases in selected Spanish regions.

Over 75% coverage of the MF cross-referencing to ICD-10-ES reached

Achievements summary:

- 95% MF (Orphacodes) cross-referencing to ICD-10-ES. Updated to the 2021 version of the Orphanet nomenclature and the third edition of the ICD-10-ES (2020).

- 5 out of the 6 regions originally involved in the RD-CODE have updated their systems to implement Orphacodes.

- 1 additional regional registry (Madrid) has been incorporated as collaborator reaching 54% coverage of the national population. Furthermore, 1 hospital (Sant Joan de Déu – Barcelona) has started to use the MF as reference for the allocation of Orphacodes at its registry.

- Expansion of the Orphacode implementation toolbox for the cases received from the population-based congenital anomalies registry by establishing cross-references among Orphacodes, ICD-10-ES and ICD-10-BPA. This approach is being piloted only for the Valencia Region as proof of concept.

Dissemination of results and recommendations:

Meetings, conferences and courses

-Reunión nacional de coordinación (National coordination meeting) – Organization and Oral presentations (April 2019) “De la RD-ACTION al RD-CODE”.

- Multi-stakeholder RD-CODE workshop – Oral presentation (June 2020) “Presentation from implementing countries on implementation progress and problems and opportunities encountered – Spain”.

- I Congreso Virtual de la Sociedad Española de Epidemiología (SEE), y de la Associação Portuguesa de Epidemiologia (APE) (1st SEE and APE virtual congress) – Oral presentation (October 2020) “Mejorando la identificación de las enfermedades raras: de CIE-10 a ORPHA”.

-16th World Congress on Public Health – Poster display (October 2020) "From ICD-10 to ORPHAcodes: paving the way towards improved identification systems for rare diseases".

- Introduction of Orphacoding in Malta – Invited lecture (November 2020) "The RD codification in Spain: RD-CODE project" and "Valencia Region (Spain) Rare Diseases and Congenital Anomalies Registries and the Orphacodes implementation".

- Internal RD-CODE workshop – Oral presentation (June 2021) "Presentation from implementing countries on achievements, challenges and next steps – Spain".

- XXXIX Reunión anual SEE - XVI Congreso APE - IX Congreso SESPAS (XXXIX SEE - XVI APE - IX SESPAS annual meeting and congress) – Oral presentation (September 2021) "La importancia de la trazabilidad en la nomenclatura Orphanet para las enfermedades raras: Comparativa 2018-2020".

- RD-CODE final workshop – Oral presentation (November 2021) "Presentation from implementing countries on key activities, conclusions and recommendations – Spain".

Publications

- Cavero-Carbonell, C., Rico, J., Echevarría-González de Garibay, L. J., García-López, M., Guardiola-Vilarroig, S., Maceda-Roldán, L. A., & Zurriaga, O. (2020). From ICD10 to ORPHAcodes: paving the way towards improved identification systems for rare diseases. European Journal of Public Health, 30 (Supplement 5).

- Rico, J., Echevarría-González de Garibay, L.J., García-López, M., Guardiola-Vilarroig, S., Maceda-Roldán, L.A., Zurriaga O., & Cavero-Carbonell. C. (2020). Mejorando la identificación de las enfermedades raras: de CIE-10 a ORPHA. Gaceta Sanitaria, 34(SC), 33.

- Zurriaga, O., Rico, J., Echevarría-González de Garibay, L.J., Guinaldo-Muñoz, J.M., Guardiola-Vilarroig, S., Maceda-Roldán, L.A., & Cavero-Carbonell, C. (2021). La importancia de la trazabilidad en la nomenclatura Orphanet para las enfermedades raras: Comparativa 2018-2020. Gaceta Sanitaria, 35(SC), 25.

- Rico, J., Echevarría-González de Garibay, L. J., García-López, M., Guardiola-Vilarroig, S., Maceda-Roldán, L. A., Zurriaga, O., & Cavero-Carbonell, C. (2021). The interoperability between the Spanish version of the International Classification of Diseases and ORPHAcodes: towards better identification of rare diseases. Orphanet Journal of Rare Diseases, 16(1), 121.

7 – Recommendations for Orphacodes implementation at the national level

The RD-CODE project has been regarded as a positive initiative in Spain in part due to the peculiarities of the National Health System (NHS). The responsibilities of the Ministry of Health are delegated to the 17 autonomous communities (AC), which have different

regulations and availability of resources. This decentralized clinical and administrative landscape hinders the recovery of RD cases at the national level, and progress on this front is often uneven. In fact, the coordination and guidance provided within the frame of the RD-CODE project has allowed the different RD regional registries to unify their criteria about the implementation of Orphacodes and to add up efforts that would otherwise have likely been futile at the national level for not being in line with those from other regions. Moreover, some regions have reported that, without the figure of a coordinator (FISABIO), a much lower number of Orphacodes would have been implemented at their registries and that the presence of such figure at the national level should continue in order to keep track of the updates and to maintain and further implement the guidelines established throughout the RD-CODE project.

Despite the fact that the figure of a coordinator (and a core working group) has shown to be beneficial for the consensual advancement of the implementation of Orphacodes, some technical aspects related to codification, or to IT development should be addressed with the support provided by specialized teams. For instance, the mapping to ICD-10-ES of particularly intricate disorders like epilepsy. The complexity and richness of both the Orphanet nomenclature and the ICD-10-ES classifications for these disorders make it quite difficult to establish cross-references without the participation of an expert coder since the terms and description rarely match between the two of them. Moreover, the coordination provided has faced a number of limitations which cannot be addressed in the short time span of the project. Such limitations are usually linked to the strategies previously defined by the regional governments in terms of RDs and the resources allocated to their RD information systems. For instance, as stated in this and previous reports, each regional registry is nourished from different information sources and, although some are common to most of them, the relative contribution to the absolute number of cases registered can be very different, especially if we compare registries with and without direct notification by clinicians. This, along with other structural differences (e.g. software, available resources, etc.), makes the use of a single strategy insufficient to attain broad implementation of Orphacodes at the national level.

One additional issue, already anticipated, that has become more evident during the RD-CODE project is that true exact mappings between Orphacodes and ICD-10-ES are an exception. Although the MF containing the cross-references to ICD-10-ES has been acknowledged as a really useful and valuable tool, it does not allow automated recodification of most of the ICD-based diagnoses collected at the Health Information Systems (HIS). Manual validation and codification of RDs with Orphacodes by clinical documentalists, hospital registries personnel or RD registries' technicians is for the time-being irreplaceable given that context. The strengthening of the structure of information professionals and their Orpha codification training should be the basis on which robust RD information systems are built. Indeed, although unclear cross-references consultation with the mapping team has been part of the workflow throughout the project, a continuous open communication channel between registries and a coordinating figure has been proposed in order to discuss and reach consensus on cross-references that may differ from the ones established in the MF (either

because of the criteria defined by the registries or because of the inclusion or modification of cross-references based on the users experience).

Overall, the use of the MF with cross-references to the classifications in use by the information sources has allowed registries to be more efficient and precise. Moreover, it is anticipated that this tool will be key for the development of a notification process from the Reference Centers, Services and Units of the National Health System (CSUR-SNS) as well as from specialized and coordination practices and general assistance for RDs (Valencia). Advances on this front would allow the use of Orphacodes directly at the information sources, which is not possible at the moment at most regional registries, avoiding the information loss that are currently experiencing due to RDs for which the ICD-10-ES is less specific. For the time-being, and given the continuous evolution of the coding systems, clear references to the ICD-10-ES and/or Orphanet nomenclature version (e.g. ICD-10-ES 2018/Orphanet 2019) used to validate or register each particular case or set of cases would be helpful to keep track of potentially outdated records and establish a regular quality-check process. This would be particularly useful for the retrieval of potential RD cases by their associated ICD-10-ES codes.

In addition, the approval of new or updated regulations at the national or regional level to drive the implementation of Orphacodes at the hospital level would represent a major step forward towards the implementation of Orphacodes beginning with the main information sources. This kind of regulations has previously proven to be pivotal for the implementation of the ICD-9-CM and the ICD-10-ES in Spain and has also favoured the advancement in the use of SNOMED-CT.

In terms of practicalities, the integration of cross-references between Orphacodes and ICD-10 based classifications in readily available browsers rather than an Excel file (e.g. incorporated to the eCIE-Maps server of the Ministry of Health) would facilitate the implementation within registries as well as provide curated up-to-date information.

Meanwhile, if a Spanish version of the MF were available, it would be much easier for the registries to incorporate the correct terms and synonyms in their systems without having to deal with manual translation or search, saving time and avoiding errors.

In conclusion, the MF cross-referencing to ICD-10-ES and the coordination efforts led by FISABIO have been much appreciated and have allowed setting a common reference database to establish datasets at the registries. However, the continuous update required and the limitations brought up during the current project to automate the identification of RDs based only on the cross-references provided by the MF, made evident that additional strategies as well as maintenance of the MF are needed:

- Establishment of a permanent coordination or working group
- Regulation-led approaches for the implementation at clinical level
- Strengthening of the information systems and networks
- Orphanet nomenclature training for clinicians, documentalists, etc...
- Inclusion of the Spanish version of the RD-CODE MF in an e-browser.

The above-listed are some of the recommendations that have been regarded as the most suitable to advance efficient implementation of Orphacodes at the national level and, consequently, for the identification of RD diagnoses.

In closing, a few lessons learned during the project that might be applicable to most implementing countries are listed below as take-home message:

- There is not a one-size-fits-all strategy to implement Orphacodes.
- Codification of RD cases with Orphacodes at the source (i.e. Clinical services) would be extremely helpful.
- Full adaptation of previously established information systems, based on a different core classification, for the use of Orphacodes is not simple and, in some instances, not even feasible.
- Cross-referencing among classifications is time and resource consuming, and requires regular maintenance. However, it bridges the gap between information sources and RD registries.

8 – List of Acronyms

AC – Autonomous Community

CIBERER – Center for Biomedical Network Research on Rare Diseases

CMBD – Minimum Basic Data Set (hospital discharges database)

CSUR-SNS – Reference Centers, Services and Units of the National Health System

ERA-EDTA – European Renal Association-European Dialysis and Transplant Association

EUROCAT – European network of population-based registries for the epidemiological surveillance of congenital anomalies

FISABIO – Foundation for the Promotion of Health and Biomedical Research in the Valencian Region

HCIS – Electronic clinical health records system

HIS – Health Information Systems

ICD-9-CM – International Classification of Diseases 9th revision – Clinical Modification

ICD-10 – International Classification of Diseases 10th revision

ICD-10-CM – International Classification of Diseases 10th revision – Clinical Modification

ICD-10-ES – Spanish International Classification of Diseases 10th revision – Clinical Modification

ICD-10-BPA – International Classification of Diseases 10th revision – British Paediatric Association

ISCIII – Carlos III Institute of Health

MF – Master file for statistical reporting with Orphacodes

NHS – National Health System

Orphacode – Orphanet nomenclature of rare diseases

RD – Rare disease

SNOMED-CT – Systematized Nomenclature of Medicine Clinical Terms

SWAN – Syndrome without a name

WP4 – Work Package 4

WP5 – Work Package 5

The links displayed in this document were last accessed on December 9 2021

9 – Authors list

Participating institutions

Foundation for the Promotion of Health and Biomedical Research in the Valencian Region (FISABIO)

Juan Rico Molíns

Berta Arribas Díaz

Clara Cavero Carbonell

Center for Biomedical Network Research on Rare Diseases (CIBERER)

Virginia Corrochano James

Ingrid Mendes

Francesc Palau

Collaborating Stakeholders

Ministry of Health of the Basque Government

Luís María Oregi Lizarralde

María del Henar Sampedro García

Mª Mercedes Ruíz-Uzkiano Viñegra

Luís Javier Echevarría González de Garibay

Public Health Office of Castile and Leon Government

María García López

José Manuel Guinaldo Muñoz

Aurora Plaza Bermejo

Rufino Álamo Sanz

Regional Department of Health of Navarre

Esther Vicente Cemborain

Elifio Cuesta Álvarez

Alberto Ovies Fernández

Eva Ardanaz Aicua

Regional Health Administration. Generalitat Valenciana

Sandra Guardiola Vilarroig

Marta Serra Briz

Enrique Llopis Mengual

Óscar Zurriaga Llorens

Murcia Regional Health Department

Joaquín A. Palomar Rodríguez

Pilar Mira Escolano

Pilar Ciller Montoya

Luis Alberto Maceda Roldán

Catalan Health Service

Xènia Acebes Roldán

Assumpta Ricart Conesa

Cristina Mallol Domínguez

Xavier Salvador Vilalta

Lluís Vilardell Ramoneda

Rare Diseases Advisor Group

Madrid Regional Health Authority, Public Health Directorate

Juan Pablo Chalco Orrego

Mª Felicitas Domínguez Berjón

Mª Dolores Esteban Vasallo

Leticia Hernaez Vicente

Ana Clara Zoni Matta

Luis Miguel Blanco Ancos (IT external service)

Sant Joan de Déu hospital (Barcelona)

Roser Francisco Bordas

Gemma Gelabert Colomé

Fernando Cid Samper