EUCERD Joint Action

MINIMUM DATA SET FOR RARE DISEASE REGISTRIES

EUCERD JOINT ACTION WP8
TASK 8.1: INTEGRATION OF REGISTRY ACTIVITIES

JANUARY 2015

Co-funded by
the Health Programme
of the European Union

Minimum Data Set

Abstract

As a crucial prerequisite for increasing the interconnectivity of data available on rare disease patients on the European level, the WP8 Task 1 team drafted a common denominator of data to be gathered on the national level and considered essential information on each patient and their disease.

This basic list of items contained in a Minimum Data Set (MDS) is intended to become the mandatory data set for all RD registries to collect. I.e., it is imagined that each RD registry must contain this basic information for each patient, so that even a very specific registry will yield comparable (e.g. epidemiologic) information that can then be used for more general purposes than just those related to the specific disease.

On a second layer, each registry could then collect additional data required and/or useful just for its specific purpose within a disease/group of diseases as well as its specific national health care context. The MDS can thus be expanded in accordance with the needs and prerequisites of any given registry. It is envisaged that this modular approach will create the greatest possible benefit for a range of stakeholders.

Methodology

The development of this MDS was mainly based on the MDS to be used by the French Banque Nationale de Données Maladies Rares (BNDMR) by condensing some of the elements contained therein and regrouping others into purpose-specific modular data sets to extend the MDS.

A further revised version was then developed in collaboration with the OSSE (Open Source Registry System for Rare Diseases) project group and has been proposed to be used as the MDS for the European RD patient registration platform.

This work was enabled by the EUCERD Joint Action (DG SANCO No. 2011 2201).

Minimal Data set - Basic Data Set - Common Data Set

This is a preliminary document, the degree of sophistication varies from high ("Basic Data Set") to low ("Leftovers").

Minimal Data Set – Basic Data Set – Common Data Set
When looking at the data that should be made available for the European Platform of Rare Disease Registries we have to use equivocal terminology:
Minimal Data Set (MDS): The smallest possible selection of items in a registry that allows to answer a single question.
Basic Data Set (BDS): The selection of items in a registry that allows to answer the standard set of questions the registry was made for.
Common Data Set (CDS): The overlap of all items of all registries that are to interoperate and which are sufficient to answer the questions planned to be addressed by the European platform.
We took the French "minimal data set" and looked how the information contained compares to other MDS. We then excluded (moved to another than basic section) some of the items, because they are not really “basic” but rather gather information required for addressing specific questions of epidemiology or resource utilization, and have added a few.

How to use this document?
1. At the time being the document shows a putative list of items for the generic "basic data set" (i.e. common data set) for all European RD registries.
2. The document will eventually help someone who wants to set-up a new registry in deciding on the type of registry and actually building a new registry according to his envisaged uses. We have therefore defined the different registry modules from which the registry will be made (ref. to table “Types of Registry”). You may use the flow chart (ref. to table “Flowchart”) for guidance.

Which type of registry do you want (basic, epidemiological, clinical, etc.)?
- Select the modules you want to include
- Within each module select items
- If necessary add items

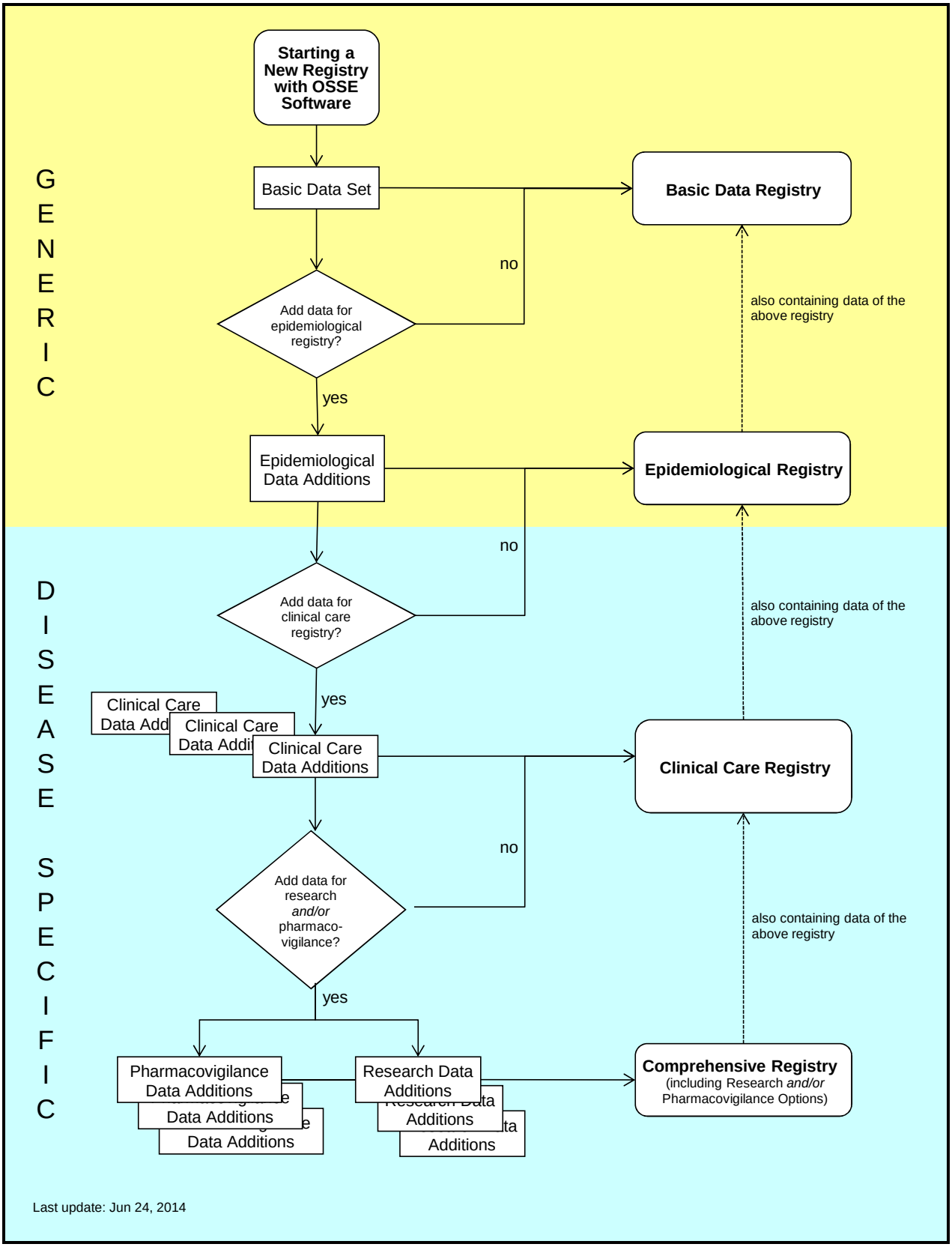
Keep in mind that:
Sorting in our tables is just a proposal and does not preclude some item to be placed somewhere else in your registry.
The forms/screens to put in data may be quite different from what you see now.
The more items you select the harder it will be to input "all" data.

Make sure you go both ways:
You want to start a registry for a specific disease and specific purpose (type of registry) and want to know which data might be necessary to meet your needs.
You have questions that you prospectively want to answer by means of a patient registry data.
In an ideal situation you should come up with the same data set whichever way you chose (two sides of the same coin).
By using this approach your registry data set will be specifically tailored to your standard questions. Since you must consider your questions will change you will have to accept that your data set may change with time.

Interoperability with other registries
For Interoperability of different registries a Common Data Set has to be defined containing a set of data that all registries should have in common. Again – in the first place - it is necessary to decide which information shall be withdrawn from the interoperating registries, then one can decide which items will make up the Common Data Set. Most likely the data contained in the Basic Data Sets will make up the common data set.

Document history:
French "minimal data set" (preliminary version); not shown.
EUCERD-JA: Minimal data set (tab version V0.1); TOF Wagner, A Pfalz; Jan 15, 2014
OSSE: Minimal Data Set (tab "Basic Items"); TOF Wagner, A Pfalz; Feb 20, 2014

Last update: Feb 26, 2014



Last update: Jun 24, 2014

Basic Data Registry
Epidemiological Registry
Clinical Care Registry
Comprehensive Care Registry - including Research Options
Comprehensive Care Registry - including Pharmacovigilance Options
Comprehensive Care Registry - including Research and Pharmacovigilance Options

BASIC

Item Group	Item No.	Item Concept	Question	Content Coding	Data Collection (one-time/repeatedly)	Comment
1. Consent	1.1	Patient consent	Does the patient give his/her consent for information to be stored in a computer data management system?	Yes	one-time (master data)	
	1.2	Patient non-opposition to the reuse of data	Has the patient been informed concening and not opposed to a future re-use of his/her de-identified data for other research purposes?	Yes No	one-time (master data)	
	1.3	Consent by legal guardian	Is the consent given by the patient's legal guardian?	Yes No	one-time (master data)	
2. Patient identification	2.1	International rare disease identifier	GUID (Global Unique Identifier)	String (automatically generated)	one-time (master data)	Whatever is available (from top to bottom) should be used
	2.2	National anonymous RD patient ID	Anonymous RD identifier	String (automatically generated)	one-time (master data)	Whatever is available (from top to bottom) should be used
	2.3	other national identifier	Patient ID	String (automatically generated)	one-time (master data)	Whatever is available (from top to bottom) should be used
		Local patient identifier	Identifier of the patient in the hospital	String	one-time (master data)	
3. Personal information	3.1	Patient's maiden name	Patient's maiden name	String	see comment	Data not part of registry but part of "Mainzelliste" (pseudonymization service)
	3.2	Patient's (married) last name	Patient's (married) last name	String	see comment	Data not part of registry but part of "Mainzelliste" (pseudonymization service)
	3.3	Patient's first name	Patient's first name as specified on the birth certificate or identity card	String	see comment	Data not part of registry but part of "Mainzelliste" (pseudonymization service)
	3.4	Patient's date of birth	Patient's date of birth as recorded on the birth certificate	Date	one-time (master data)	Consider using only MM/YYYY
	3.5	Gender	Patient's gender	Female Male Undetermined Unknown (for the foetus)	one-time (master data)	
		Country of birth	Patient's country of birth	Country code	one-time (master data)	
	3.7	Country of residence	Patient's country of residence	Country code	one-time (master data)	
	3.6	Foetus (if applicable)	Whether information is recorded for a foetus	Yes No	one-time (master data)	

BASIC

Item Group	Item No.	Item Concept	Question	Content Coding	Data Collection (one-time/repeatedly)	Comment
7. Disease history	7.1	Age at onset	Age at which symptoms first appeared	Antenatal At birth XX year(s) and XX month(s) Undetermined	one-time (master data)	
	7.2	Age at diagnosis	Age at which the diagnosis was made	Antenatal At birth XX year(s) and XX month(s) Not available	one-time (master data)	Not applicable (e.g. no diagnosis yet)
8. Diagnosis	8.2	Diagnosis of the rare disease	Diagnosis retained by the RD center	Orphanet code	one-time (master data)	
	8.1	Current state of the diagnosis	What is the current status of the diagnosis?	Ongoing Likely Confirmed Undetermined	one-time (master data)	
5. Vital status	5.1	Patient's vital status	Is the patient still alive?	Yes No	repeatedly	Update of the data base once per year
	5.2	Patient's date of death	Patient's date of death	Date	one-time (master data)	If 5.1 is answered with "no" this item appears
		Death due to the rare disease	Is the death due to the rare disease the patient is suffering from?	Yes No Unknown	one-time (master data)	If 5.1 is answered with "no" this item appears
		If death is not linked to the rare disease, main course of death	If death is not due to the rare disease, what is the main cause of death?	ICD-10 code	one-time (master data)	If 5.1 is answered with "no" this item appears

EPIDEMIOLOGICAL

Item Group	Item No.	Item Concept	Question	Content Coding	Data Collection (one-time/repeatedly)	Comment
		Genetics	Gene identified	Yes No Partly	one-time	
8. Diagnosis	8.4	Sproadic or familial disease	Is the case isolated or familial at the time of observation (assessed by the RD professional)	Sporadic Familial	one-time	
3.		City of birth	Patient's city of birth	City code	one-time	
		City of residence	Patient's city of residence	City code	one-time	
4. Family information	4.1	Patient born from a relationship between related parties	Is the patient born from a relationship between related parties?	Yes No Unknown	one-time	
		Propositus	First patient recorded in a rare disease center for a given family	Yes No	one-time	
		Relationship with the propositus	Describes the relationship between a patient of the family and the propositus	Brother Sister […] Maternal grandfather Maternal grandmother Unknown	one-time	This question only appears if the person is not the propositus

CLINICAL CARE

Item Group	Item No.	Item Concept	Question	Content Coding	Data Collection (one-time/repeatedly)	Comment
6. Care pathway		Patient addressed by	The patient may have been referred to by a health care professional. This assesses the attractiveness of the RD reference center for professionals.	Come by himself Patient association GP Paediatrician (non-hospital practice) Paediatrician (hospital practice) Center for maternal and child health Geneticist Gynecologist/obstetrician Other specialist (city/hospital) Center of support (CAMSP, CPMM, SESSAD, ...) Multidisciplinary center for prenatal diagnosis Center for prenatal screening Other		
	6.1	Patient's date of inclusion in the RD center	Date at which the patient was recorded in the RD center	Date		
Care activities		Date of care performed by the patient	Date at which the declared patient care was performed	Date		
		Context of the care activity	Context indicating the type of care performed (useful information for the RD refernce center annual report	Consultation Multidisciplinary consultation Day hospital Traditional hospital Advice on medical record during consultation Expert advice on medical record Advice given in another ward Teleconsultation Other context		
		Objectives of the care activity	Objectives of the declared care activity	Diagnosis Implementation of support Follow-up Genetic counseling Prenatal diagnosis Preimplantation diagnostics Emergency support Medical act Research protocol		
		Occupation of the personnel performing the care activity	Type of personnel performing the declared activity	Psychologist Physical therapist Genetic counselor Nurse Speech therapist Specialist teacher Physician Other professional		
		Personnel performing the activity	Name of the personnel performing the activity	String		

CLINICAL CARE

Item Group	Item No.	Item Concept	Question	Content Coding	Data Collection (one-time/repeatedly)	Comment
Patient Status		Weight	Patient's current weight (kg)	Numeric	repeatedly	
		Height	Patient's current height (cm)	Numeric	repeatedly	
		Disabilities	How is the disability according to [add disease specific indicator]		repeatedly	
		QoL	Patient's quality of life according to [add suitable indicator]		repeatedly	
8. Diagnosis	8.3	Additional or unusual signs associated with the rare disease	Additional or unusual signs associated with the rare disease diagnosis (assesed by the RD professional)	HPO code		
9. Treatment	9.1	Is a specific treatment for the rare disease in process?	Is a specific treatment for the rare disease in process? (N.b.: A definition of what constitutes a disease-specific treatment is necessary. The so-called "comfort" treatments are not considered here.)	Yes No		This is somewhat un-precise: what is a specific treatment? We need a definition:like treatment that requires specific know-how which usually is only available in a Centre of expertise?
		Ongoing treatment for the rare disease	Name of the ongoing specific treatment for the rare disease. Only rare disease treatments are considered here.	Orphanet / ATC code		
10. Pre- and neo-natal		Medically assisted procreation	Has a medical assistance program for procreation been used?	Yes No		this is a speccific scientific question
	10.1	Presence of pre-natal malformation	Does the patient have a prenatal malformation?	No - LOINC Unique Multiple		
	10.2	Full-term (if else, clarify the term)	Was the patient born at full-term? If not, clarify the term.	Yes No		
	10.3	Height at birth	Patient's height at birth (cm)	Numeric		
	10.4	Weight at birth	Patient's weight at birth (g)	Numeric		
	10.5	Head circumference at birth	Patient's head circumference (cm)	Numeric		
		Foetopathology	Has a foetopathological examination been performed?	Yes (with or without autopsy) No		

RESEARCH

Item Group	Item No.	Item Concept	Question	Content Coding	Data Collection (one-time/repeatedly)	Comment
11. Research		Patient participating in a protocol	Is the patient currently involved in a research protocol (cohort, ...)?	Yes No		
	11.1	Agreement to be contacted for a protocol	Does the patient give permission to be contacted for a research protocol?	Yes No		
	11.2	Patient having previously given a biological sample for research	Has the patient already given a bilogical sample for research?	Yes No		does this need separation from 11.3?
	11.3	Patient having previously given a biological sample for molecular diagnosis	Has the patient already given a biological sample for molecular diagnosis?	Yes No		
		Link to a biobank				

PHARMACOVIGILANCE

Item Group	Item No.	Item Concept	Question	Content Coding	Data Collection (one-time/repeatedly)	Comment

LEFTOVERS

BASIC DATA

Item Group	Item No.	Item Concept	Question	Content Coding	Data Collection (one-time/repeatedly)	Comment
7.		Assessment of the diagnosis at center admission	Was the patient's diagnosis appropriate when admitted to the RD center?	Missing Inappropriate Appropriate		
Diagnosis confirmation		Confirmation method of the diagnosis	What method was used to confirm the diagnosis?	Clinical Molecular genetic Cytogenetic Biochemical Biological Imaging Other		
		Biological method used to assess the diagnosis (if applicable)	Specify the diagnosis method used	Chromosomal (karotype, FISH) Array-CGH Targeted sequencing (Sanger) High throughput sequencing (targeted, exome, genome) Other methods		
		Mutation (if applicable)	What is (are) the mutation(s) involved?	Nomenclature taking into account the individual's genotype (Mutnomen: www.hgvs.org/mutnomen)		
		Penetrance (if applicable for a seemingly healthy patient)	Is the present subject (i.e. the subject of this collection of information), who carries the mutation, sick or not?	Yes No Unknown		

Data Set for Rare Disease Patient Registries Recommended for European Cooperation * (Version 0.1)

Item Group	Item No.	Item Concept	Question	Content Coding	Data Collection (one-time/repeatedly)	Comment
1. Consent	1.1	Patient consent	Does the patient give his/her consent for information to be stored in a computer data management system?	Yes		
	1.2	Patient non-opposition to the reuse of data	Has the patient been informed concening and not opposed to a future re-use of his/her de-identified data for other research purposes?	Yes No		
	1.3	Consent by legal guardian	Is the consent given by the patient's legal guardian?	Yes No		
2. Patient identification	2.1	International rare disease identifier	GUID (Global Unique Identifier)	String (automatically generated)		Whatever is available (from top to bottom) should be used
	2.2	National anonymous RD patient ID	Anonymous RD identifier	String (automatically generated)		Whatever is available (from top to bottom) should be used
	2.3	other national identifier	Patient ID	String (automatically generated)		Whatever is available (from top to bottom) should be used
		Local patient identifier	Identifier of the patient in the hospital	String		
3. Personal information	3.1	Patient's maiden name	Patient's maiden name	String	one-time	
	3.2	Patient's (married) last name	Patient's (married) last name	String	repeatedly	
	3.3	Patient's first name	Patient's first name as specified on the birth certificate or identity card	String	one-time	
	3.4	Patient's date of birth	Patient's date of birth as recorded on the birth certificate	Date	one-time	
	3.5	Gender	Patient's gender	Female Male Undetermined Unknown (for the foetus)	one-time	
		City of birth	Patient's city of birth	City code		
		Country of birth	Patient's country of birth	Country code		
		City of residence	Patient's city of residence	City code		
	3.6	Foetus (if applicable)	Whether information is recorded for a foetus	Yes No		
	3.7	Country of residence	Patient's country of residence	Country code	repeatedly	

Item Group	Item No.	Item Concept	Question	Content Coding	Data Collection (one-time/repeatedly)	Comment
4. Family information		Propositus	First patient recorded in a rare disease center for a given family	Yes No		
		Relationship with the propositus	Describes the relationship between a patient of the family and the propositus	Brother Sister […] Maternal grandfather Maternal grandmother		
	4.1	Patient born from a relationship between related parties	Is the patient born from a relationship between related parties?	Yes No Unknown		
5. Vital status	5.1	Patient's vital status	Is the patient still alive?	Yes No	repeatedly	Update of the data base at least once per year?
	5.2	Patient's date of death	Patient's date of death	Date	one-time	Update of the data base at least once per year?
		Death due to the rare disease	Is the death due to the rare disease the patient is suffering from?	Yes No Unknown		
		If death is not linked to the rare disease, main course of death	If death is not due to the rare disease, what is the main cause of death?	ICD-10 code		
6. Care pathway		Patient addressed by	The patient may have been referred to by a health care professional. This assesses the attractiveness of the RD reference center for professionals.	Come by himself Patient association GP Paediatrician (non-hospital practice) Paediatrician (hospital practice) Center for maternal and child health Geneticist Gynecologist/obstetrician Other specialist (city/hospital) Center of support (CAMSP, CPMM, SESSAD, ...) Multidisciplinary center for prenatal diagnosis Center for prenatal screening Other		
	6.1	Patient's date of inclusion in the RD center	Date at which the patient was recorded in the RD center	Date		

Item Group	Item No.	Item Concept	Question	Content Coding	Data Collection (one-time/repeatedly)	Comment
Care activities		Date of care performed by the patient	Date at which the declared patient care was performed	Date		
		Context of the care activity	Context indicating the type of care performed (useful information for the RD refernce center annual report	Consultation Multidisciplinary consultation Day hospital Traditional hospital Advice on medical record during consultation Expert advice on medical record Advice given in another ward Teleconsultation Other context		
		Objectives of the care activity	Objectives of the declared care activity	Diagnosis Implementation of support Follow-up Genetic counseling Prenatal diagnosis Preimplantation diagnostics Emergency support Medical act Research protocol		
		Occupation of the personnel performing the care activity	Type of personnel performing the declared activity	Psychologist Physical therapist Genetic counselor Nurse Speech therapist Specialist teacher Physician Other professional		
		Personnel performing the activity	Name of the personnel performing the activity	String		
7. Disease history	7.1	Age at onset	Age at which symptoms first appeared	Antenatal At birth XX year(s) and XX month(s) Undetermined		
		Assessment of the diagnosis at center admission	Was the patient's diagnosis appropriate when admitted to the RD center?	Missing Inappropriate Appropriate		
	7.2	Age at diagnosis	Age at which the diagnosis was made	Antenatal At birth XX year(s) and XX month(s) Not available		Not applicable (e.g. no diagnosis yet)

Item Group	Item No.	Item Concept	Question	Content Coding	Data Collection (one-time/repeatedly)	Comment
8. Diagnosis	8.1	Current state of the diagnosis	What is the current status of the diagnosis?	Ongoing Likely Confirmed Undetermined	repeatedly	
	8.2	Diagnosis of the rare disease	Diagnosis retained by the RD center	Orphanet code		
	8.3	Additional or unusual signs associated with the rare disease	Additional or unusual signs associated with the rare disease diagnosis (assesed by the RD professional)	HPO code		
	8.4	Sproadic or familial disease	Is the case isolated or familial at the time of observation (assessed by the RD professional)	Sporadic Familial		
Diagnosis confirmation		Confirmation method of the diagnosis	What method was used to confirm the diagnosis?	Clinical Molecular genetic Cytogenetic Biochemical Biological Imaging Other		
		Biological method used to assess the diagnosis (if applicable)	Specify the diagnosis method used	Chromosomal (karotype, FISH) Array-CGH Targeted sequencing (Sanger) High throughput sequencing (targeted, exome, genome) Other methods		
		Mutation (if applicable)	What is (are) the mutation(s) involved?	Nomenclature taking into account the individual's genotype (Mutnomen: www.hgvs.org/mutnomen)		
		Penetrance (if applicable for a seemingly healthy patient)	Is the present subject (i.e. the subject of this collection of information), who carries the mutation, sick or not?	Yes No Unknown		
9. Treatment	9.1	Is a specific treatment for the rare disease in process?	Is a specific treatment for the rare disease in process? (N.b.: A definition of what constitutes a disease-specific treatment is necessary. The so-called "comfort" treatments are not considered here.)	Yes No		This is somewhat un-precise: what is a specific treatment? We need a definition:like treatment that requires specific know-how which usually is only available in a Centre of expertise?
		Ongoing treatment for the rare disease	Name of the ongoing specific treatment for the rare disease. Only rare disease treatments are considered here.	Orphanet / ATC code		

Item Group	Item No.	Item Concept	Question	Content Coding	Data Collection (one-time/repeatedly)	Comment
10. Pre- and neo-natal		Medically assisted procreation	Has a medical assistance program for procreation been used?	Yes No		this is a speccific scientific question
	10.1	Presence of pre-natal malformation	Does the patient have a prenatal malformation?	No - LOINC Unique Multiple		
	10.2	Full-term (if else, clarify the term)	Was the patient born at full-term? If not, clarify the term.	Yes No		
	10.3	Height at birth	Patient's height at birth (cm)	Numeric	one-time	
	10.4	Weight at birth	Patient's weight at birth (g)	Numeric	one-time	
	10.5	Head circumference at birth	Patient's head circumference (cm)	Numeric	one-time	
		Foetopathology	Has a foetopathological examination been performed?	Yes (with or without autopsy) No		
11. Research		Patient participating in a protocol	Is the patient currently involved in a research protocol (cohort, ...)?	Yes No		
	11.1	Agreement to be contacted for a protocol	Does the patient give permission to be contacted for a research protocol?	Yes No		
	11.2	Patient having previously given a biological sample for research	Has the patient already given a bilogical sample for research?	Yes No		does this need separation from 11.3?
	11.3	Patient having previously given a biological sample for molecular diagnosis	Has the patient already given a biological sample for molecular diagnosis?	Yes No		
		Link to a biobank				

** This table is based on the French National Rare Diseases Minimum Data Set, RD MDS v1.08*