REPORT OF
WORK PACKAGE 1
EVALUATION OF THE NUMBER OF PATIENTS WITH RARE DISEASES

Lead partner of Work Package	**Maastricht University**
WP n° and title	WP1 Evaluation of offer and demand in the field of rare diseases in the EMR
Activity n° and title	A1.1. Need and demand of EMR rare disease patients and social and health(care) professionals
Dissemination level	Public
Delivery Date	24 January 2020
Status	Final
Language in which the report is available	**EN – FR – NL - DE**

With the support of

TABLE OF CONTENTS

AUTHORS' LIST

Name/Surname	Institution name	Contact e-mail
Leading Author(s)		
Rok Hrzic	Maastricht University	r.hrzic@maastrichtuniversity.nl
Co-Author(s)		
Timo Clemens	Maastricht University	timo.clemens@maastrichtuniversity.nl
Peter Schröder-Bäck	Maastricht University	peter.schroder@maastrichtuniversity.nl
Helmut Brand	Maastricht University	helmut.brand@maastrichtuniversity.nl

We would also like to acknowledge Joyce Loridan from Solidaris and Clara Noirhomme from ANMC for their invaluable assistance with extracting insurance claims data for the Belgian EMR.

The EMRaDi project was carried out under the Interreg V-A Euregio Meuse-Rhine programme and was supported by the European Union and the European Regional Development Fund for an amount of EUR 1,687,675.

The Interreg V-A Euregio Meuse-Rhine (EMR) programme invested almost EUR 100 million in the development of the Interreg-region until 2020. This area stretches out from Leuven in the west to the borders of Cologne in the east, and runs from Eindhoven in the north all the way down to the border of Luxemburg. Over 5.5 million people live in this cross-border region, where the best of three countries merges into a truly European culture.

With the investment of EU funds in Interreg projects, the European Union directly invests in the economic development, innovation, territorial development and social inclusion and education of this region.

With the support of

PROJECT DESCRIPTION

"EMRaDi" stands for **Euregio Meuse-Rhine Rare Diseases**.
The project started on 1st October 2016 and ended on 31st March 2020

This project involved a **cross-border cooperation** between health insurers, university hospitals, patient associations and a university in the Euregio Meuse-Rhine. It was part of the European Union INTERREG V-A Euregio Meuse-Rhine programme.

Thanks to their long experience in cross-border healthcare, the project partners have decided to join forces in the specific field of rare diseases. This EMRaDi project was innovative in the sense that it was a patient-oriented and cross-sectoral project. The consortium of partners included the major health players who support rare disease patients and their relatives in their day-to-day rare disease patient pathway.

Through **the project activities**, the EMRaDi project aimed to:

- increase the transparency of needs and availability of services in the field of rare diseases in the Euregio Meuse-Rhine (EMR);
- develop EMR models for *rare disease patient pathways* in order to draw up patient-oriented recommendations in synergy with national and European developments;
- improve the network of healthcare providers, health insurance providers and patient organisations and raise (public) awareness of rare diseases.

The general long-term aim was to **improve the quality of life of these patients**.

www.emradi.eu

LEGAL ISSUES

This report was produced within the framework of the EMRaDi project. The facts and views expressed in this publication are the sole responsibility of the authors and do not necessarily reflect the position of the EMRaDi partner organisations.

The methodology utilized in this report was examined by the METC azM/UM and declared not to pose a threat to individual health or privacy under Dutch law (non-WMO declaration 2018-0927).

EXECUTIVE SUMMARY

Work package 1 aimed to characterize the need and demand of Euregio Meuse-Rhine (EMR) rare disease (RD) patients, and report on the number of identified patients with rare diseases and their specific rare diseases in the region (activity T1.1). In particular, disease burden estimates for a list of 60 rare diseases were sought.

We utilised a cross-sectional secondary analysis approach to electronic health records from hospital information systems of large regional university hospitals and relevant databases of insurance claims data. The institutions included are Hospitalier Universitaire de Liège (CHU), Universitätsklinikum Aachen, Maastricht Universitair Medisch Centrum Plus (MUMC+), Solidaris, Mutualité chrétienne (MC) and Mutualités Libres – Onafhankelijke Ziekenfondsen (MLOZ). The German Institute for Medical Documentation and Information (DIMDI) and Vektis, which hold the German and Dutch health insurance claims database, respectively, were contacted and requests for data were submitted, but they did not provide the data by December 2019. Due to the sensitive nature of the data, we asked the data holders to collate the individual data into per-disease aggregate datasets in-house before submitting them to Maastricht University for further analysis. The rare disease patients were identified using a combination of diagnosis and regional identifiers. We operationalised the "burden" of rare diseases in the EMR as the number of distinct patients and the total number of patient visits per year associated with selected rare diseases, as well as their associated costs.

Due to lack of standardisation of the different information systems in terms of resources and governance structures, it was not possible to ensure a full alignment in terms of the years included and type of data extracted. Triangulated data was available for eight rare diseases: Chronic myeloid leukaemia (CML), Duchenne muscular dystrophy, Galactosemia type 1, Huntington disease, Phenylketonuria (PKU), Polycythaemia Vera (PV), Rett syndrome, and Silver-Russell syndrome. The cost data was shared only by Belgian data holders.

We can note substantial variation between the burdens various diseases represent, as well as between the three hospitals that contributed the data and between Belgian hospital and insurance claims data. The data also exhibit marked temporal heterogeneity (variations in annual prevalence of 2-3-fold within a single institution and disease). On the whole, the included diseases represented tens to hundreds of hospital visits per year on average. Similar variation between diseases was present in the cost data with haematological diseases tending to be the costliest, which is likely the result of the availability of specialised pharmacological treatment. In most of rare diseases for which data cost was available, the majority of spending occurred outside specialised hospital care in either primary care, supportive care, or in other costs.

It is recommended that the European Union, national and regional government, and individual healthcare providers strengthen the capacity for access to RD patient data for research purposes – particularly research into provision of health services for RDs. We also recommend an expanded and systematic survey of the burden of rare diseases that goes beyond those included in this study. Both of these aims can be supported by future actions in the EMR.

1 INTRODUCTION

Work package 1 (WP1) aimed to characterize the need and demand of Euregio Meuse-Rhine (EMR) rare disease (RD) patients, specifically to answer the question: what is the number of identified patients with rare diseases and their specific rare diseases in the region? Another EMRaDi report provides an overview of the complex needs of rare disease patients that have been described in the scientific literature[1]. The present report will tackle the question of the burden of rare diseases in the EMR as defined by the number of RD patients that seek medical care in the region and the associated healthcare costs.

The European Commission defines diseases as rare if their prevalence is lower than 1 in 2000. An estimated 6000-8000 diseases meet this criterion. Because such a vast scope of inquiry is untenable for any single research project, the EMRaDi partners created short and long lists of diseases to focus on. The short list contains eight representative rare diseases: Chronic myeloid leukaemia (CML), Duchenne muscular dystrophy, Galactosemia type 1, Huntington disease, Phenylketonuria (PKU), Polycythaemia Vera (PV), Rett syndrome, and Silver-Russell syndrome. The long list contains 60 rare diseases (Supplementary Table 1). The process of selecting these diseases is reported in the EMRaDi Final Report (Appendix 5).

Due to the difficulty of coding rare diseases in health information systems and the estimated high prevalence of undiagnosed rare diseases, estimating rare disease prevalence remains extremely challenging. The most commonly cited estimated prevalence rate of rare diseases is 6-8% of the population. This translates into 240,000 to 320,000 patients in the EMR. However, recent hospital- and registry-based studies in Italy[2] and Hong Kong[3] found prevalence rates much smaller (0.3% and 1.5%, respectively), which revises the possible lower end of EMR patient numbers to between 12,000 and 60,000 patients.

Prevalence estimates are an important input in planning medical care and social support services. In the context of the EMRaDi project, EMR RD prevalence estimates for the selected eight diseases will inform the recommendations regarding RD care pathways in the region, as well as improvements regarding reimbursement procedures. A lower prevalence would allow us to consider more individualized treatment of RD patients in the region. Conversely, a higher prevalence restricts this option and requires a careful consideration of feasibility of any service provision on a larger scale.

The remainder of this Report is structured as follows: The methods describe our approach to estimating EMR RD patient numbers and associated costs of treatment. The results section provides an overview of the number of patients diagnosed with the eight diseases that recently sought treatment in the region. In the final section, we provide our reflection on the process of data retrieval and offer some lessons learned and recommendations for improving the RD data infrastructure in the EMR, identify what our findings mean for developing better RD care pathways, and describe the strengths and limitations of this study.

2 METHODS

2.1 RESEARCH DESIGN

We utilised a cross-sectional secondary analysis approach to electronic health records from hospital information systems of large regional university hospitals and relevant databases of insurance claims data.

2.2 DATA SOURCES AND ACCESS PROCEDURE

The sources of data were the electronic health records of three university hospital centres that were partners in the EMRaDi project, namely Centre Hospitalier Universitaire de Liège (CHU), Universitätsklinikum Aachen, and Maastricht Universitair Medisch Centrum Plus (MUMC+).

The planned sources for health insurance claims data in Germany and the Netherlands were the German Institute for Medical Documentation and Information – DIMDI, which holds the German health insurance claims database (DaTrav database) – and Vektis, which holds the Dutch health insurance claims database. However, neither of these sources provided the results in the two-year time window available for their data to be included in the analysis.

As of writing this Report (January 2020), the database of the Agence InterMutualiste (AIM) which collects all the data to carry out specific studies related to health care in Belgium does not include data allowing the identification of patients according to their diagnosis and therefore patients with rare diseases. Instead, we worked with three Belgian health mutuals to access the relevant data – Solidaris and Mutualité chrétienne (MC), also partners in the EMRaDi project – and Mutualités Libres – Onafhankelijke Ziekenfondsen (MLOZ), another Belgian health mutual and associated partner.

2.3 DATA COLLECTION PROCESS

For each of the institutions listed above (data holders), we followed the established procedures to request research access, where these existed (DIMDI and Vektis). Where such procedures did not exist, we worked with the relevant departments of the data holders (e.g., medical information or billing departments) to establish a route to access.

Figure 1 summarises the data collection process, which took place over the 2017-2019 period.

Due to the sensitive nature of the data, we asked the data holders to collate the individual data into per-disease aggregate datasets in-house before submitting them to Maastricht University for further analysis. This way we minimized any potential risks of data leaks during the transfer of the data and protected the anonymity of the patients. This was also crucial to ensure buy-in from the relevant data holders and respect the provisions of the General Data Protection Regulation (GDPR) that had recently come into force.

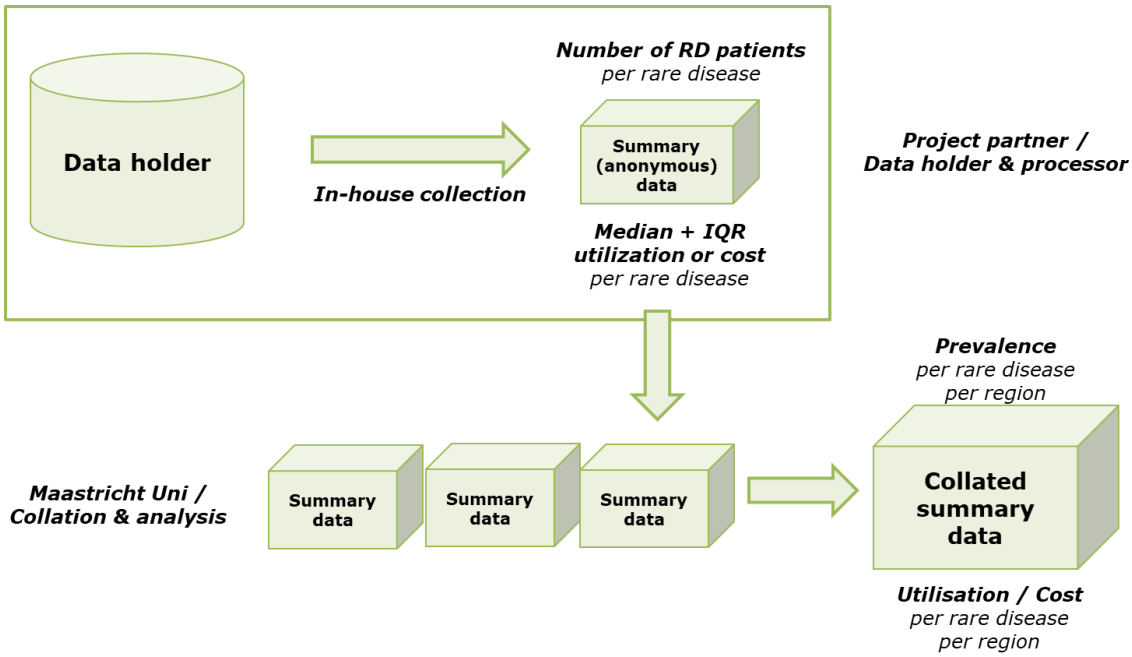

Figure 1 Data collection and analysis procedure

2.4 SAMPLE

Our initial plan was to extract prevalence and cost information on all 60 diseases from the long list (Supplementary Table 1). However, given the lack of established procedures for research access to the data and limited staff resources at the university hospital centres, we instead focused on the shortlist of eight rare diseases there, and pursued the long list with the data holders of the health insurance claims only. The rare disease patients were identified using a combination of diagnosis and regional identifiers. Since it was the data holders extracting the data, they implemented the sampling strategy according to the guidelines described below.

ICD-10 or ICD-9 codes were used (depending on the year extracted and centre performing the extraction) to identify the relevant individuals. The translation between the two coding systems was performed by hand searching the ICD-9 code book and extracting relevant diagnoses and their codes (see Table 1). For insurance claims data, the ICD code-based identification strategy was supplemented with additional keywords and drug prescription data (protocol in French available by request).

For regional identification, we assumed that the catchment area of the three university hospital centres corresponded to the EMR geographic area. This assumption was made in order to simplify the identification of patients with these data holders, not only in light of their resources, but also because it directly represents the rare disease patient burden that these institutions are faced with. With health insurance claim data holders, postcodes were used to identify the population living in the EMR.

Table 1. Short list of rare diseases and the relevant ICD-9 and ICD-10 codes

Disease	ICD-9 code (annual average for EMR)	ICD-10 code (annual average for EMR)
Chronic myeloid leukaemia (CML)	205.1x	C92.1x
Duchenne muscular dystrophy	359.1x	G71.0x
Galactosemia type 1	271.1x	E74.21
Huntington disease	333.4x	G10x
Phenylketonuria (PKU)	270.1x	E70.0x E70.1x
Polycythaemia vera (PV)	238.4x	D45x
Rett syndrome	330.8x	F84.2x
Silver-Russell syndrome	759.8x	Q87.1x

2.5 OUTCOMES

We operationalised the "burden" of rare diseases in the EMR as the number of distinct patients and the total number of patient visits per year associated with selected rare diseases, as well as their associated costs. We made efforts to break the costs down into several categories, particularly costs associated with the labour of medical professionals and costs of medicines.

2.6 LEGAL BASIS AND DATA SECURITY

The methodology utilized in this report was examined by the METC azM/UM and declared not to pose a threat to individual health or privacy under Dutch law (non-WMO declaration 2018-0927). We also obtained the consent of data holders.

Data security was established and maintained by conducting the data extraction on-site, which was performed by the data holders themselves. There were no off-site data transfers involving individual patient data. Each data holder securely transferred the collated data (i.e. table of patient counts per disease and year) to Maastricht University, where it was held on internal servers. Access to the hospital-specific results datasets was limited to the first author of this Report.

3 RESULTS

3.1 SCOPE OF DATA COLLECTED

Table 2 summarises the availability of data over time per data holder; the shaded areas represent available data. Due to lack of standardisation of the different information systems in terms of resources and governance structures, it was not possible to ensure a full alignment in terms of the time included and data extracted. No overall alignment in terms of time could be achieved, therefore the average annual prevalence numbers are reported. The cost data was shared only by Belgian stakeholders.

Table 2. Distribution of collected data by time and place

Year	Liege	Aachen	Maastricht	BE insurers
2013				*
2014	*			*
2015				*
2016				*
2017	*			
2018				

*** Cost data available**

3.2 THE BURDEN OF SELECTED RARE DISEASES IN THE EMR

Table 3 summarises the number of patients and total hospital visits related to selected rare diseases in the EMR. The overall situation is represented by the numbers of distinct patients and hospital visits, which were calculated as the averages over time summed over the three hospitals. The geographic heterogeneity is represented by the range, which was calculated as the minimal and maximal average values over time reported.

Table 3. The number of distinct patients and total hospital visits related to selected rare diseases in the EMR

Disease	N of distinct patients[1]	(range[2])	N of hospital visits[1]	(range[2])
Chronic myeloid leukaemia (CML)	156	(42-65)	649	(60-429)
Duchenne muscular dystrophy	144	(6-90)	290	(8-188)
Galactosemia type 1	24	(0-24)	52	(0-52)
Huntington disease	132	(4-74)	215	(5-137)
Phenylketonuria (PKU)	35	(2-30)	85	(1-80)
Polycythaemia vera (PV)	187	(34-99)	605	(50-322)
Rett syndrome	47	(0-45)	91	(0-90)
Silver-Russell syndrome	83	(21-37)	134	(34-53)

[1] Values are averages over time summed over the three hospitals, rounded to closest integer.
[2] Values are minimal and maximal averages over time, rounded to closest integer.

Substantial variations can be seen between the burdens various diseases represent, as well as between the three hospitals that contributed the data. The former is most likely the reflection of the wide variation in natural prevalence of various rare diseases, ranging from ultra-rare to more "common" rare diseases. The latter is most likely a reflection of hospitals focusing on different diseases. The differences in implementing the sampling strategy also likely contributed to the variation. We consider it unlikely that the geographic heterogeneity was caused by differences in the natural prevalence of these diseases.

The data also exhibit marked temporal heterogeneity (2-3-fold variations in annual prevalence within a single hospital and disease; not shown in this Report), but no clear increasing or decreasing trend.

3.3 DATA COMPARISON FOR HOSPITAL AND INSURANCE CLAIMS PREVALENCE IN THE BELGIAN EMR

As we were only able to acquire data from both hospital and insurance claims data sources in Belgium, Table 4 compares the number of hospital visits related to selected rare diseases for the Belgian EMR. This comparison allows for a rudimentary triangulation of values and a relative appraisal of accuracy of reported values.

Table 4. The number of hospital visits related to selected rare diseases in the Belgian EMR

Disease	N of hospital visits (hospital data)	N of hospital visits (insurance claims data)
Chronic myeloid leukaemia (CML)	60	184
Duchenne muscular dystrophy	8	15
Galactosemia type 1	0	/
Huntington disease	5	39
Phenylketonuria (PKU)	4	7
Polycythaemia vera (PV)	50	25
Rett syndrome	0	8
Silver-Russell syndrome	48	/
Note: All values are averages over time, rounded to closest integer.		

We can note a substantial discrepancy between data based on hospital information systems and data based on insurance claims data. There are several potential reasons for this that would require further exploration, including:

- Imperfect transfer of data between hospital and insurance information systems;
- Difference in diagnosis coding for clinical and reimbursement purposes;
- Differences in catchment area (there is another large university hospital centre in the region that did not participate in this study);
- Differences in identification method (within this study), and
- Trends over time.

However, there is substantial agreement between the data sources regarding relative importance of the various disease burdens, i.e., both identify CML as the most frequently encountered rare disease from our list.

3.4 THE COST OF SELECTED RARE DISEASES IN THE BELGIAN EMR

Because we were only able to acquire cost data from both hospital and insurance claims data sources in Belgium, Table 5 provides insight into the average annual treatment costs for selected rare diseases in the Belgian EMR.

Table 5. The average annual cost of treatment of selected rare diseases in the Belgian EMR

Disease	Average annual cost of treatment per patient (hospital data)	Average annual cost of treatment per patient (insurance claims data)
Chronic myeloid leukaemia (CML)	12,826 EUR	22,762 EUR
Duchenne muscular dystrophy	5,111 EUR	13,795 EUR
Galactosemia type 1	/	/
Huntington disease	4,477 EUR	17,495 EUR
Phenylketonuria (PKU)	4,251 EUR	6,322 EUR
Polycythaemia vera (PV)	10,396 EUR	60,305 EUR
Rett syndrome	/	12,020 EUR
Silver-Russell syndrome	11,729 EUR	/
Note: All values are averages over time, rounded to the closest integer. Both covered and out-of-pocket expenses are included.		

As expected, hospital costs are smaller, as they tend to exclude the costs associated with ambulatory care. Other reasons for the discrepancy mentioned in the previous section may also play a large role. However, as in the case of prevalence data, there is good agreement on the relative position of cost of different rare diseases.

We can observe that haematological diseases tend to be the most expensive, which is likely the result of the availability of specialised pharmacological treatment. According to the insurance claims data available (not reported), pharmaceuticals represent more than 50% of the healthcare-related costs of CML and PV. The same goes for PKU.

Insofar as neurological and developmental disorders are concerned, the insurance claims data suggest that nursing and other costs are substantial and can exceed half of the total treatment cost.

The cost of hospital care for all 35 rare diseases for which insurance claims data were available (Supplementary Table 1) represents between 1 and 69% of total treatment related costs. In half of the diseases the cost of hospital care represented less than one-third of all costs.

4 CONCLUSIONS

4.1 KEY FINDINGS

Our study successfully collected tangible information on the state of the information systems that underpins real-world data research into rare diseases in the EMR, and on the burden of selected rare diseases in the EMR.

In terms of the information system, our experience was that it is highly fragmented and currently not conducive to further health services research. This experience is driven by three main factors:

1. The lack of interoperability of hospital information systems across national borders;
2. The lack of a clear pathway to access relevant data for the purpose of health services research;
3. The lack of adequate human resources of the data holders that could support robust health services research by providing information on data available in their systems and/or timely access to this data or data summaries.

We are grateful to the information specialists for the different data holders that we came in contact with, because it was apparent that supporting our study was beyond their usual tasks. It also leads us to conclude that real-world data research within the context of rare diseases is something that does not systematically occur with these data holders and therefore requires further support and development in the future.

With regards the burden of selected rare diseases, we found substantial variation between diseases, between data holders within the same disease, and over time within the same disease and data holder. This leads us to conclude that one-size-fits-all approaches to optimising rare disease care will not be appropriate. However, we also found that rare disease diagnoses give rise to tens or hundreds of hospital visits each year. This means that *ad hoc* unsystematic approaches to organising patient care that rely on the administrative and managerial skills of individual medical experts are insufficient.

4.2 RECOMMENDATIONS

It is recommended that efforts be made by the European Union, national and regional governments, and individual healthcare providers, to strengthen the capacity for access to RD patient data for research purposes – particularly research into the provision of health services for RDs.

This will require the development of common (and cross-border) guidelines on collecting, storing, and accessing the relevant medical information, and supporting the implementation of these processes with adequate human and other resources. The EMR could play a leading role in this process as an example EU region, particularly given that EMRaDi already brought relevant stakeholders into close and long-lasting contact.

In light of the uncovered heterogeneity in disease burdens that the different rare diseases represent, we would also recommend an expanded and more systematic survey of the burden of rare diseases that goes beyond our selection of eight and sixty. A clustering exercise on healthcare utilisation data could highlight valuable similarities between biologically different diseases in terms of patient needs

and provider burdens. Such insights could reduce the number of distinct RD patient pathways that need to be designed, piloted, and implemented.

In line with the previous two areas of recommendations, the funders are asked to consider a more permanent regional (EMR) intelligence effort dedicated to systematic and robust cross-border RD data collection and analysis that would focus on information related to the provision of health services. Such an effort would effectively complement the existing registries focused on genetic, biomedical, and clinical information and could improve the quality of life of RD patients by supporting more efficient access to therapies that already exist (even if they are not curative).

4.3 STRENGTHS AND LIMITATIONS OF THE STUDY

By combining RD patient data across six data holders from three countries, this study provides the first regional intelligence on the burden of RD in the EMR in terms of patient numbers, hospital visits, and costs. It also provides insight into existing barriers for future health services research related to RDs in the region. The recommendations provide potential future directions for data governance and research that could tangibly improve the RD patient's healthcare experience with in the mid- to long-term.

The main limitation of this study is its distributed execution of sampling, which highly likely biased the results, and thus made reliable head-to-head comparisons between the data holders less likely. It is also likely that our sampling identified patients, hospital visits, and costs that had little or nothing to do with the selected rare diseases, and that it failed to identify some that did. We therefore avoided discussing our results in a comparative sense in the conclusions or interpreting them in a precise manner, but rather focused on what the magnitude of the effects and relative sizes of the disease burdens of selected RDs mean for improving the provision of care for RD patients in the EMR.

5 LIST OF ABBREVIATION AND ACRONYMS

(in alphabetic order)

CML	Chronic myeloid leukaemia (CML)
EMR	Euregio Meuse-Rhine
ICD	International Classification of Diseases
PKU	Phenylketonuria
PV	Polycythaemia vera
RD	Rare disease

6 BIBLIOGRAPHY

1 Hrzic R, Clemens T, Schroder-Back P, Brand H. Report on the demands and needs of patients. 2018. EMRaDi project, www.emradi.eu

2 Mazzucato M, Dalla Pozza LV, Manea S, Minichiello C, Facchin P. A population-based registry as a source of health indicators for rare diseases: the ten-year experience of the Veneto Region's rare diseases registry. *Orphanet J Rare Dis* 2014;9:37.

3 Chiu ATG, Chung CCY, Wong WHS, Lee SL, Chung BHY. Healthcare burden of rare diseases in Hong Kong – adopting ORPHAcodes in ICD-10 based healthcare administrative datasets. *Orphanet J Rare Dis* 2018;13:147.

7 SUPPLEMENTARY MATERIAL

Supplementary table 1. The long and short lists of rare diseases included in the project.

No.	indication area(s)	RD Subgroup	Important examples	ICD 10 Code
1	*Neurological*	*Ataxia/Coordination disorders*	*Friedreich Ataxia*	*G11.1*
2	Neurological	Ataxia/Coordination disorders	Spinocerebellar ataxias (Type 41-43): Example Type 1	G11.8
3	***Neurological***	***Rare genetic neurodegenerative disease***	***Huntington disease***	***G10***
4	*Neurological*	*Rare genetic neurodegenerative disease*	*Hereditary Spastic Paraparesis*	*G11.4*
5	*Neurological*	*Rare genetic neurodegenerative disease*	*Amyotrophic lateral sclerosis*	*G12.2*
6	*Neurological*	*Rare neurodegenerative disease*	*Progressive supranuclear palsy*	*G23.1*
7	Neurological	Myopathies	None provided	None provided
8	Neurological	Neuropathies	Sodium channelopathy-related small fibre neuropathy	G62.8 G60.8
9	*Neurological*	*Epilepsy syndromes*	*epileptic encephalopathy* *(West syndrome (G40.4) and the Landau-Kleffner syndrome (G80.3)*	*G40.3*
10	*Neurological*	*Phakomatoses*	*Neurofibromatosis Type 1*	*Q85.0*
11	Neurological	Phakomatoses	Neurofibromatosis Type 2	Q85.0
12	***Neurological***	***Myotonic dystrophies***	***Duchenne muscular dystrophy***	***G71.0***
13	Neurological	Myotonic dystrophies	Becker muscular dystrophy	G71.0
14	*Neurological*	*Myotonic dystrophies*	*Type 1: Steinert myotonic dystrophy*	*G71.1*
15	Neurological	Myotonic dystrophies	Type 2: Proximal myopathy, PROMM disease	G71.1
16	*Neurological*	*Hereditary neuropathy*	*Charcot-Marie-Tooth disease*	*G60.0*
17	*Neurological*	*Rare neuromuscular disease*	*Myasthenia gravis*	*G70.0*

18	*Neurological*	*Frontotemporal Dementia*	*Not provided*	*G31.0*
19	Neurological	Rare neurodegenerative disease	Early-onset Parkinson Syndrome	G20
20	Neurological	Not provided	Generalised Primary Dystonia	ex G24.1
21	***Haematological***	***Myeloproliferative neoplasms (MPN)***	***Bcr-Abl positive: Chronic myeloid leukaemia (CML)***	***C92.1***
22	***Haematological***	***Myeloproliferative neoplasms (MPN)***	***Bcr-Abl negative: Polycythaemia vera (PV) or Vaquez disease***	***D45***
23	*Haematological*	*Myeloproliferative neoplasms (MPN)*	*Essential thrombocythemia (ET)*	*D47.3*
24	*Haematological*	*Myeloproliferative neoplasms (MPN)*	*Primary myelofibrosis (PMF)*	*D47.4*
25	Haematological	Myeloproliferative neoplasms (MPN)	Chronic eosinophilic leukaemia (CEL)	D47.5
26	Haematological	Myeloproliferative neoplasms (MPN)	Chronic neutrophilic leukaemia (CNL)	D47.1
27	Haematological	Myeloproliferative neoplasms (MPN)	MPN-unclassifiable (MPN-U)	Not provided
28	Haematological	Myeloproliferative neoplasms (MPN)	Eosinophilic MPN with rearrangement of PDGFRA, PDGFRB, or FGFR1	Not provided
29	*Haematological*	*Bone marrow failure syndromes*	*Aplastic anaemia (AA)*	*D61 (D61.0-3;8-9)*
30	*Haematological*	*Bone marrow failure syndromes*	*Paroxysmal nocturnal haemoglobinuria (PNH)*	*D59.5*
31	Haematological	Mastocytosis	Systemic mastocytosis	D47
32	*Haematological*	*(Hypoplastic) myelodysplastic syndrome (MDS)*	*Not provided*	*D46*
33	*Haematological*	*Hereditary haemolytic disease*	*Sickle cell anaemia, Drepanocytosis*	*D57.0 D57.1 D57.2*
34	*Haematological*	*Rare genetic haematological disorder*	*Haemophilia*	*Not provided*
35	Haematological	A plasma cell disorder	AL amyloidosis	E85.9

36	Syndromic	Rare genetic syndromic intellectual disability	Kabuki syndrome	Q87.0
37	***Syndromic***	***Rare genetic syndromic intellectual disability***	***Rett syndrome***	***F84.2***
38	Syndromic	Rare Genetic Syndromes, Imprinting Disorders (Q87)	Not provided	Q87
39	**Syndromic**	**Rare Genetic Syndromes, Imprinting Disorders (Q87)**	**Silver-Russell syndrome**	**Q87.1**
40	Syndromic	Rare Genetic Syndromes, Imprinting Disorders (Q87)	Temple-Baraitser Syndrome	Q87.1
41	Syndromic	Rare Genetic Syndromes, Imprinting Disorders (Q87)	Beckwith-Wiedemann syndrome	Q87.3
42	*Syndromic*	*Rare Genetic Syndromes, Imprinting Disorders (Q87)*	*22q11.2 deletion syndrome*	*D82.1*
43	*Syndromic*	*Rare (cyto)genetic syndromes with congenital abnormalities and/or intellectual disability*	*Prader-Willi syndrome*	*Q87.1*
44	*Syndromic*	*Rare genetic disease, chromosomal anomalies*	*Turner syndrome*	*Q96.0 Q96.1 Q96.2 Q96.3 Q96.4 Q96.8 Q96.9*
45	*Syndromic*	*Short stature, typical facial dysmorphism and congenital heart defects*	*Noonan syndrome*	*Q87.1*
46	**Metabolic**	**Galactosemia**	**Galactosemia (type 1)**	**E74.2**
47	***Metabolic***	***inborn error of amino acid metabolism***	***Phenylketonuria***	***E70.0 E70.1***
48	Metabolic	renal tubular amino acid transport disorder	Cystinuria	E72.0
49	*Metabolic*	*lysosomal storage disease*	*Fabry disease*	*E75.2*
50	Metabolic	lysosomal storage disease	Hunter disease, Mucopolysaccharidosis type 2 (MPS2))	E76.1
51	Metabolic	lysosomal storage disorder	Gaucher disease	E75.2

52	Metabolic	Glycogen storage disease	Pompe disease, Glycogen storage disease due to acid maltase deficiency	E74.0
53	*Metabolic*	*Very rare inherited multisystemic disease due to excessive copper deposition*	*Wilson disease*	*E83.0*
54	*Other*	*Systemic disease of connective tissue*	*Marfan Syndrome*	*Q87.4*
55	*Other*	*Nonneoplastic pulmonary disease*	*Idiopathic pulmonary fibrosis*	*J84.1*
56	*Other*	*Rare bone disease*	*Osteogenesis imperfecta*	*Q78.0*
57	*Other*	*Generalized disorder of small arteries, microvessels and connective tissue*	*Systemic sclerosis*	*M34.0 M34.1 M34.2 M34.8 M34.9*
58	*Other*	*Rare birth defects, rare neural tube defects*	*Isolated spina bifida*	*Q05.0 Q05.1 Q05.2 Q05.3 Q05.4 Q05.5 Q05.6 Q05.7 Q05.8 Q05.9*
59	*Other*	*Rare genetic disorder, channelopathy*	*Cystic fibrosis, Mucoviscidosis*	*E84.0 E84.1 E84.8 E84.9*
60	Other	Rare genetic disorders, rare hepatic and respiratory disorders	Alpha-1-antitrypsin deficiency	Not provided

Note: The diseases on the short list are listed here in bold font. Insurance claims data was available for italicised items.