

EJP RD

European Joint Programme on Rare Diseases

H2020-SC1-2018-Single-Stage-RTD

SC1-BHC-04-2018

Rare Disease European Joint Programme Cofund



Grant agreement number 825575

Del 10.5

**Report on the State of the Art of existing
resources in Europe**

Organisation name of lead beneficiary for this deliverable:

Partner 1 – INSERM

Contributors: EMBL-EBI, RaDiCo

Due date of deliverable: month 09

Dissemination level: Public

Table of contents

1	Background	3
1.1	Resources fragmentation	3
1.2	Researcher's needs	3
2	Objectives	5
3	Methods	5
3.1	Working group creation	5
3.2	Identification of infrastructural components necessary for research for diagnostic improvement and therapeutic development	6
3.3	Benchmarking and prioritisation of existing resources.....	6
3.3.1	2017 survey on RD resources	7
3.3.2	2019 survey on RD resources	7
4	Results	9
	Resources assessment	11
5	Conclusions	18
ANNEX I.	List of research resources for rare diseases.....	20
ANNEX II.	List of WG-identified resources contacted for the 2017 survey	44
ANNEX III.	List of questions asked in the 2017 survey	47
ANNEX IV.	List of questions asked in the 2019 survey	49
ANNEX V.	Results of the 2017 survey on RD resources	50
ANNEX VI.	Results of the 2019 survey on RD resources	52
ANNEX VII.	List of members of the 2017 working group.....	58

List of figures

Figure 1.	Survey ERNs: Is this resource of utmost importance in your research? (ERNs survey 2019)	4
Figure 2.	RD research resources landscape (see ANNEX I for resources details).	8
Figure 3.	Expected effect of EJP RD on research efforts.....	19

List of tables

Table 1.	Evaluation of resources based on the 2017 and 2019 surveys on resources..	17
-----------------	---	-----------

1 Background

1.1 Resources fragmentation

Resources relevant for rare disease (RD) research already in existence, such as knowledge bases, databases, registries, tools, and infrastructures, are significant. However, despite the multitude of efforts, most of them allowed by European funding, in providing the RD research community with appropriate tools and platforms, research advancements in the RD field remain hampered by the duplications of efforts and fragmentation of resources making their efficient use challenging¹.

The aim of collaborative work in Pillar 2 is geared towards decreasing fragmentation and maximizing European capacity to enable better and more efficient research on RD by bringing the interdisciplinary key players, their assets and know-how together with the end-users. The foreseen EJP RD Virtual Platform aims at decreasing fragmentation and overlaps by building a single access point to find and query resources, and to provide data analysis facilities. Not only will it be made possible by coordinated access to the most important catalogues for RD (e.g. Orphanet, RD-Connect biobanks and patient registry catalogues) and tools (customized ELIXIR bio.tools), as well as to the main data deposition and analysis resources for genomics, phenomics and multi-omics data, but access to infrastructures providing support to clinical and translational research will be improved. This will be ultimately accessible to the whole RD community.

1.2 Researcher's needs

The priority of the EJP RD is that needs of researchers are met but these needs must be identified first. In 2019, the EJP RD Pillar 2 sent a survey to European Reference Networks (ERNs) in order to identify the RD researchers' needs (WP10.2.2; Deliverable 10.1). Two hundred and ninety-one responses were received (response rate of 31%) that allowed Pillar 2 to characterize current RD research as follows:

1. Almost all (97%) ERN units do clinical research, and many do also translational and basic research.
2. The main purposes of research are (in order of importance): (a) improve diagnostics (73%), (b) identify disease modifiers (69%), (c) develop novel therapies (64%), and (d) develop disease models (22%).

¹ Thompson, R., Johnston, L., Taruscio, D., Monaco, L., Bérout, C., Gut, I. G., et al. (2014). RD-Connect: an integrated platform connecting databases, registries, biobanks and clinical bioinformatics for rare disease research. *Journal of General Internal Medicine*, 29 Suppl 3(S3), S780–7. <http://doi.org/10.1007/s11606-014-2908-8>

3. Researchers use mostly the following data types: (a) patient information (95%), (b) natural history (91%), (c) biological samples (85%), (d) electronic medical health records (81%), (e) phenotypes (73%), (f) medical images (71%).
4. Out of 165 respondents, 79% of researchers do not deposit their -omics data in open or controlled access resources (esp. due to legal jurisdiction and complexity of use issues).
5. ICD, OMIM, ORDO, and HPO are the main ontologies/standards used to annotate data.
6. Very few researchers (13%) have FAIRified (<1%) or are in the process of FAIRifying data (12%), due in part to insufficient understanding of FAIRification as well as technical barriers.
7. Services of utmost importance for research are: (a) support to conduct clinical trials (69%), (b) facility to find patient registries (63%), (c) support to conduct translational research (53%), (d) facility to find biobanks/biosamples/cell lines (45%), (e) facility to find software and tools (40%), and (f) databases of re-usable data related to RDs (-omics) (30%).
8. Research infrastructures and resources of utmost importance for research are: (a) international/national registries (84% of respondents), Orphanet catalogue of RD resources (61% of respondents), ERDRI (50% of respondents), RaDiCo (26% of respondents), RD-Connect sample and registry and biobank finder (17% of respondents), BBMRI-ERIC directory of biobanks (15% of respondents).
9. ERNs often do not know about existing research infrastructures or resources (see Figure 1).

This survey made clear that researchers need resources to be better exposed, ordered, and accessible to save them time, money, workforce, and reduce overall data management and analysis hassles.

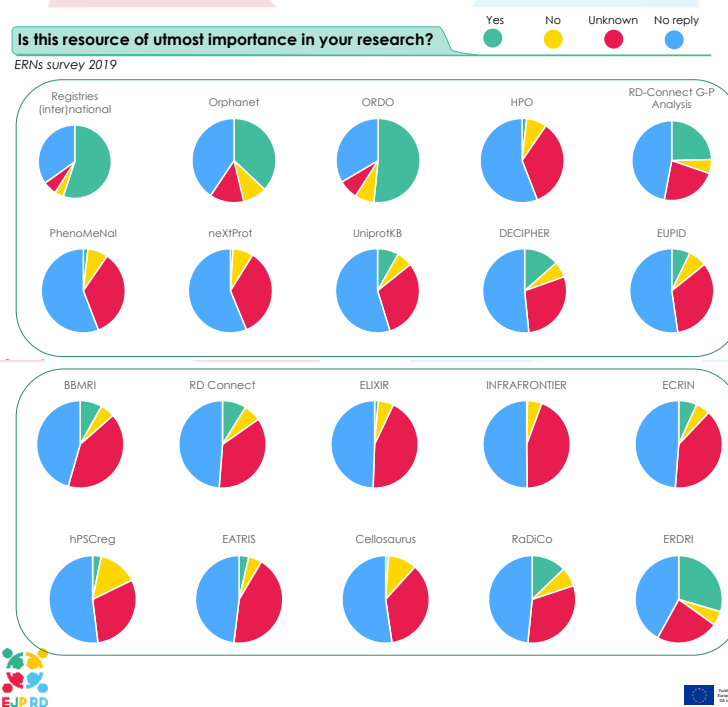


Figure 1. Survey ERNs: Is this resource of utmost importance in your research? (ERNs survey 2019). (See ANNEX I for resources details.)

2 Objectives

The aim of this report is threefold: (1) to provide an overview of existing resources that can be valuable for RD research, (2) to assess them in terms of readiness for this use, and (3) to estimate the value that the planned EJPRD Virtual Platform can add to these resources by making them usable cooperatively. This report is not intended to expose the Pillar 2 strategy, which has been detailed in the deliverable 10.1. It focuses on infrastructural RD research resources such as platforms, tools, and standards. Resources of expertise, such as ERNs, and data sources such as individual registries were considered out of scope for this deliverable and therefore excluded. These resources can still be found in the list of research resources for RD provided in ANNEX I.

3 Methods

3.1 Working group creation

The overall concept of the Pillar 2 is to improve the research cycle from bedside to bench and back again. That is the aim of the planned EJP RD Virtual Platform (VP) to be created within the Pillar. Three steps were taken in order to ensure the viability of the VP and the representativeness in terms of RD-research resources:

- (1) Mapping the needs of researchers: identification of the types of resources needed at each step of both the diagnostics and therapeutics research pathways.
- (2) Identification of existing RD-research resources: identification all the available resources that match the needs listed in step (1).
- (3) Description of existing RD-research resources: description of each resource in order to detect eventual overlaps and gaps in the imagined bed-to-bench-to bedside path.

In order to complete each of these steps, a working group (WG) composed of 46 people (ANNEX VII), representing researchers, ERNs, RD-specific and non-RD-specific resources, patient organization representatives and member state representatives, was created in 2017 to establish a list as complete as possible of known initiatives/resources useful for RD research. The WG classified these resources according to their usefulness towards the two main purposes, i.e. diagnostics or therapeutics. Members of the WG were also to make sure that, to the best of their knowledge, all types of resources needed to foster diagnosis and therapy development were represented.

The WG was to:

- (1) Define what was needed to achieve the goal (e.g. foster diagnosis- oriented research, provide interoperable clinical data and tools, ontologies, imaging

data, biobanks, genome/phenome data repositories, match-making facilities and knowledge bases).

- (2) Agree on criteria for assessment of each resource compared to the needs: readiness or potential/limits.
- (3) Conduct a survey
- (4) Identify overlaps/duplications and gaps.
- (5) Distribute the analysis workload.

3.2 Identification of infrastructural components necessary for research for diagnostic improvement and therapeutic development

Eighty-two resources were identified by the WG (incl. 27 [33%] RD-specific, see ANNEX I) and considered necessary to fasten RD research.

Among these 82 identified resources, 22 were considered to be important for the diagnostics research pathway, 24 for the therapeutics research pathway, and 36 for both the diagnostics and the therapeutics research pathways.

This list is not exhaustive (e.g. patients' organisations such as EURORDIS are not represented although they provide valuable information to researchers) but includes all well-known and largely used platforms, tools, and standards for rare diseases research.

3.3 Benchmarking and prioritisation of existing resources

Considering the large number of research resources currently available, it was necessary, after their identification and categorization, to assess how optimized these resources are for RD as well as select a few to be used for benchmarking in the EJP RD Virtual Platform.

3.3.1 2017 survey on RD resources

In 2017, from the 82 identified by the WG, 40 European RD resources were selected and contacted (ANNEX II). These resources were asked to fill out a survey in order to assess their current level of optimization for RD research. The list included both diagnosis and therapy-oriented resources. It also covered a large spectrum of research needs including pharmacology and molecular screening, phenotype/clinical features mining, genomic variant interpretation, proteomics, clinical trials support, catalogues of biobanks and registries, patients cohorts as well as data sharing and access standards. Out of the 40 contacted resources, 20 responses were received.

Resources were to answer questions (list of questions in ANNEX III) on their infrastructure's characteristics, i.e. maturity of the project, involvement in RD, plans for further/increased involvement in RD, interoperability, transparency of terms of use/IP issues, sustainability, users, involvement of patient organisation, collaboration or plans with ERNs, input to EJP RD, and output from EJP RD.

3.3.2 2019 survey on RD resources

In 2019, a second list of 20 RD resources was assessed. This list focused on resources for depositing, integrating and storing quality controlled (meta)data including registries (JRC-ERDRI, RD- Connect Registry and Biobank Finder, known also as IDcards), patient cohorts (RaDiCo, BBMRI-ERIC Directory), biobanks (RD-Connect Sample Catalogue, RD-Connect Registry and Biobank Finder, BBMRI-ERIC Directory), cell lines (hPSCreg, Cellosaurus), mouse models (INFRAFRONTIER), raw omics data (Pride, MetaboLights) and genome-phenome platforms (DECIPHER, RD-Connect GPAP).

These resources were also contacted and asked to answer questions (list of questions in ANNEX IV) on their infrastructure's characteristics, i.e. involvement in RD, kind of data dealt with, standards and vocabularies used, access model, FAIR status of the resource, projects/initiatives the resource is involved in, input to EJP RD, and output from EJP RD.

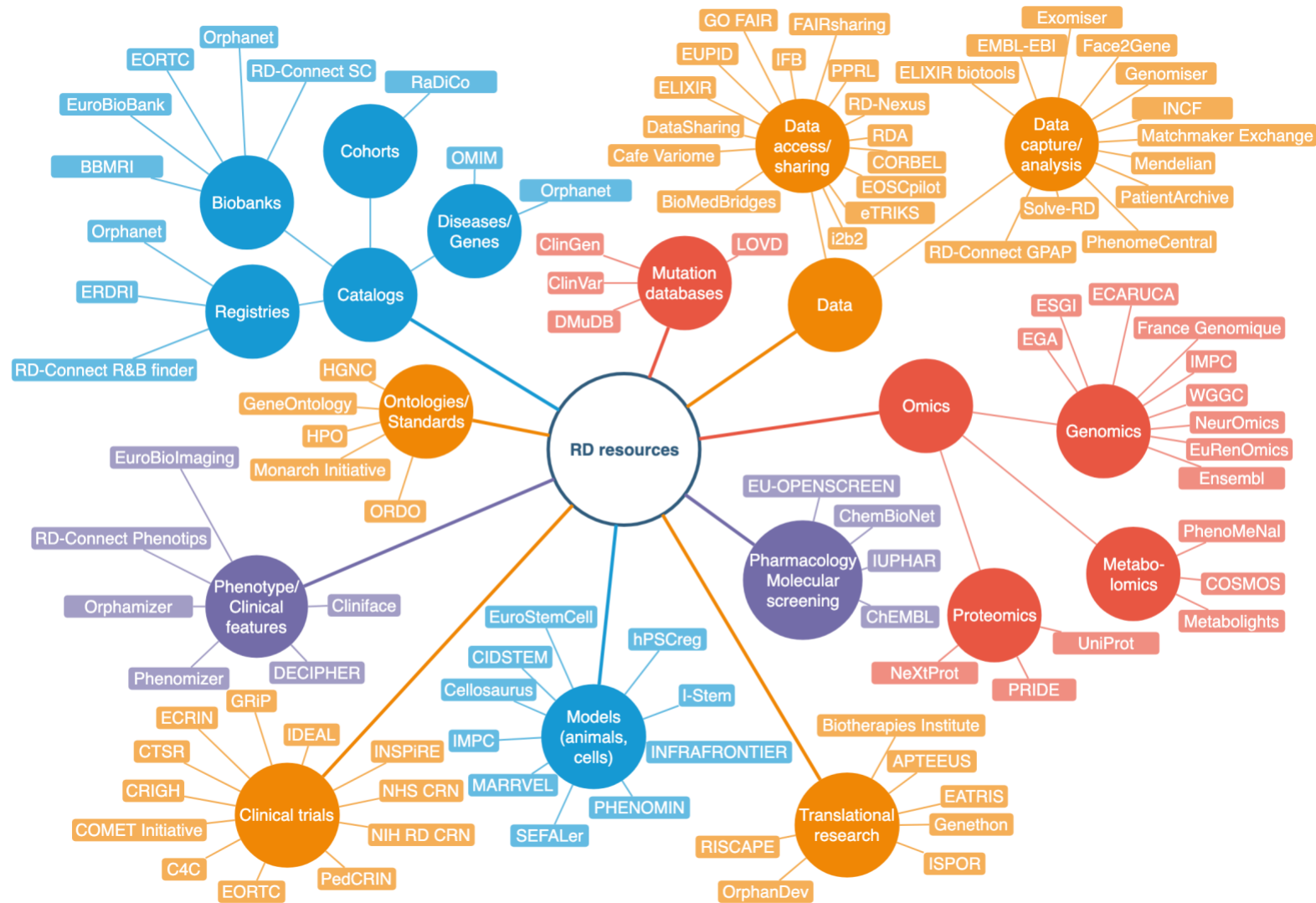


Figure 2. RD research resources landscape (see ANNEX I for resources details).

4 Results

Results of the 2017 and 2019 surveys are shown in ANNEX V and ANNEX VI respectively. Questions of the 2017 and 2019 surveys are listed in ANNEX III and ANNEX IV respectively.

The final 2017 survey was composed of the following 20 resources: ChemBioNet, DECIPHER, EATRIS-ERIC, EORTC-SPECTA, EU-OPENSSCREEN, Euro-BioImaging (EuBI), EuroBioBank, ECRIN and PEDCRIN, European Platform on Rare Diseases Registration (EU RD Platform), FAIR data services (GO-FAIR), FAIRsharing (formerly BioSharing), I-Stem, IDEAL, Orphanet and derived (Orphadata and ORDO), PPRL, RaDiCo, RD-Connect, SOLVE-RD, and UniProt.

The final 2019 survey was composed of the following 20 resources: BBMRI-ERIC, Cellosaurus, DECIPHER, Ensembl Variant Effect Predictor (VEP), EUPID, EuroBioBank, FAIR evaluator, HPO, hPSCreg, INFRAFRONTIER, MetaboLights, neXtProt, Orphanet (Orphadata, ORDO, HOOM), PRIDE, RaDiCo, RD-Connect GPAP, RD-Connect Sample Catalogue, RD-NEXUS, EU RD Platform, and UniProt.

The following six resources were in common between the 2017 and 2019 surveys: DECIPHER, EU RD Platform, EuroBioBank, Orphanet (ORDO, Orphadata), RaDiCo, and UniProt.

In total, both surveys interrogated the following 35 resources: BBMRI-ERIC, Cellosaurus, ChemBioNet, DECIPHER, EATRIS-ERIC, Ensembl VEP, EORTC-SPECTA, EU-OPENSSCREEN, EUPID, EuBI, EuroBioBank, ECRIN, EU RD Platform, FAIR evaluator, FAIRsharing, GO-FAIR, HPO, hPSCreg, I-Stem, IDEAL, INFRAFRONTIER, MetaboLights, neXtProt, Orphanet, ORDO, PEDCRIN, PPRL, PRIDE, RaDiCo, RD-Connect, RD-Connect GPAP, RD-Connect Sample Catalogue, RD-NEXUS, SOLVE-RD, and UniProt.

From the results of both the 2017 and 2019 surveys, the “readiness for RD research” of the 35 resources, i.e. how well optimized is the resource for RD research, was evaluated (Table 1). To assess this level of optimization, 13 criteria were evaluated and separated into two blocks:

Core criteria (fundamental criteria that must be fulfilled to ensure a minimum level of optimization and usability of the resource for RD research):

- RD standards: Are the standards, ontologies, and vocabularies used by the resource relevant in the RD context?
- Open access: Is it open access?
- FAIR resource/API: Is the resource FAIR or declared FAIR and/or does it have accessible application programming interfaces (APIs)?
- Interoperability: Do the tools/results ensure semantic and/or technical interoperability?
- RD-specific: Is the resource specific for the RD community or does it conduct activities specifically related to the RD field?

Additional criteria (criteria that, if fulfilled, increase significantly the resource's relevance and usability for the RD community):

- External quality assessment: Is the resource IRDiRC-recognized^{1,2} and/or is it an ELIXIR Core Data Resources³?
- FAIR-sharing-listed resource: Is the resource recommended by FAIR-sharing⁴?
- Sustainability: Is the project sustainable (further funding envisioned at the end of the project)?
- Collaborations: Are there collaboration or plan to collaborate with other partners, especially ERNs, and/or participate in other projects?
- Maturity: Have the goals of the project/resource been achieved and are tools delivered fully operational?
- Transparency/IP issues: Are licensing and intellectual property issues/terms and conditions for re-use of tools and/or results clearly established?

¹ Ayme, S. (2016). IRDiRC-recommended. *European Journal of Human Genetics*, 24(7), 955–955. <http://doi.org/10.1038/ejhg.2015.236>

² Lochmüller, H., Le Cam, Y., Jonker, A. H., Lau, L. P., Baynam, G., Kaufmann, P., et al. (2017). “IRDiRC Recognized Resources”: a new mechanism to support scientists to conduct efficient, high-quality research for rare diseases. *European Journal of Human Genetics*, 25(2), 162–165. <http://doi.org/10.1038/ejhg.2016.137>

³ Durinx, C., McEntyre, J., Appel, R., Apweiler, R., Barlow, M., Blomberg, N., et al. (2016). Identifying ELIXIR Core Data Resources. *F1000Research*, 5, 2422. <http://doi.org/10.12688/f1000research.9656.2>

⁴ FAIRsharing website: <https://fairsharing.org>

- Future RD involvement (plans to adapt to RD): Are there plans to develop (additional) RD-specific tools/methods in the future or is the resource involved in the development of new therapies for RD?
- Involvement of patients' organisations (PO): Are patients or patient organisations involved in the project?

Resources assessment

See all results of the evaluation in Table 1.

- 80-100% of core criteria fulfilled (four or five "+" in core criteria):

Resources: EuroBioBank, EU RD Platform, Orphanet, RD-Connect, RD-Connect Sample catalogue, BBMRI-ERIC, RaDiCo, ORDO, HPO, RD-Connect GPAP, DECIPHER, Ensembl, MetaboLights, and hPSCreg.

Catalogues:

EuroBioBank is a long-standing network of rare disease biobanks in Europe and well optimized for RD research. Now operating in its 17th year, EuroBioBank agreed to be the de facto biobank of RD-Connect in 2016.

The JRC's EU-RD Platform' main function is to provide interoperability for RD data collection and data sharing. The Platform is developing tools to support existing registries and the creation of new registries, EU-level standards for data collection and data sharing, recommendations for data collection and data sharing, as well as training for new and existing registries.

RaDiCo, Orphanet, and RD-Connect (incl. RD-Connect Sample Catalogue) are long-standing initiatives specific to RD and as such are very well optimized for RD research. Although some sustainability issues still need to be resolved for all three resources, they remain a staple within the RD community. Orphanet and RD-Connect GPAP are IRDiRC-recognized resources. Orphadata (derived from Orphanet) is an ELIXIR Core Data resource.

BBMRI-ERIC is neither RD-specific nor an IRDiRC-recognized resource, however, their repository can be easily searched to access RD-specific biobanks. Although several criteria could not be evaluated in these surveys, BBMRI is a European research infrastructure consortium fully sustainable and involved in many other projects including RD-specific ones. It is also listed in the FAIR-sharing's list of resources even if not officially recommended.

Ontologies:

ORDO is provided and maintained by Orphanet. It provides a structured, machine readable vocabulary for rare diseases derived from the Orphanet knowledge base. ORDO contributes to improving the interoperability of data on rare diseases across the globe and across the fields of health care and research. ORDO was first released in 2014 and is now integrated in several bioinformatics projects and infrastructures around the world in order to improve diagnosis and treatment. HOOM, for HPO-ORDO Ontological Module, was released in 2018 and qualifies annotations between clinical entities (ORDO) and phenotypic abnormalities (HPO). ORDO is an IRDiRC-recognized resource and, as part of Orphadata, an Elixir Core Data Resource.

HPO is a broadly-used and a well optimized resource for RD research. It provides a standardized vocabulary of phenotypic abnormalities encountered in human disease. The HPO is currently being developed using the medical literature, Orphanet, DECIPHER, and OMIM. HPO is an IRDiRC-recognized resource.

Data sharing/analysis:

RD-Connect GPAP is an RD-specific tool to collect and analyse phenotypic and genomic patients' data. It uses searchable unique identifiers, data is accessible by APIs and metadata uses standard ontologies, including HPO, OMIM, ORDO, and HGNC. RD-Connect GPAP is an IRDiRC-recognized resource.

The DECIPHER platform and database is developed and maintained by Genome Research Limited (operating as the Wellcome Trust Sanger Institute), a company incorporated in England and registered as a charity. DECIPHER's main aim is to curate the clinical genome and to identify the regions of the genome that are important for health and disease. DECIPHER is an RD-specific resource that ensures that all its bioinformatics tools and features remain current and that users are able to utilize the most appropriate and up-to-date reference databases available on the World Wide Web. DECIPHER is an IRDiRC-recognized resource.

Omics:

Ensembl is a powerful toolset for the analysis, annotation, and prioritisation of genomic variants and is also a well-developed project that is well optimised for RD research. It is actively developed to provide enhanced annotation functionality in rare disease studies. Ensembl uses ORDO as one of its standards, can be openly and freely accessed, and has REST APIs available.

MetaboLights provides the scientific community with a comprehensive, high-quality and freely accessible resource for Metabolomics data and tools. MetaboLights deals with metadata associated with studies including human clinical data as well as raw data from techniques (e.g. Mass spectrometry or NMR). MetaboLights is an open-access FAIRsharing-recommended resource that has accessible APIs for uploading to the repository.

Models:

hPSCreg was launched in 2007 by the European Commission as a means to provide an overview of hPSC lines available for research. It has since grown to a worldwide continuously growing registry of currently over 2600 hPSC lines from 30 countries. Data on the cell lines is collected in a standard form to enable data consistency and comparison of lines from different sources (i.e. countries and research institutes). A standard unique identifier is assigned to every registered line to enable unambiguous re-identification in research outputs.

- 40-60% of core criteria fulfilled (two to three "+" in core criteria):

Resources: EORTC-SPECTA, ECRIN, PEDCRIN, INFRAFRONTIER, RD-NEXUS, SOLVE-RD, neXtProt, PRIDE, and UniProt.

Clinical and translational research:

EORTC core element criteria could not be evaluated at this time. However, EORTC is a well-established, fully sustainable, and RD-involved non-profit research organisation. EORTC was created in 1962 and its activities are currently continuing under the Screening Patients for Efficient Clinical Trial Access (SPECTA) program. EORTC holds a strong track record in the field of RD and is expanding its network of clinical sites with the ERN EURACAN healthcare professionals and is working towards full integration with EURACAN and SPECTA serving as research infrastructure for EURACAN. EORTC is active in the development of new therapies for RD. It is also fully interoperable and using appropriate standards for all data recordings. Patients' organisations are systematically involved in EORTC studies.

ECRIN was created in 2013 and is a sustainable, not-for-profit, distributed infrastructure with the legal status of a European Research Infrastructure Consortium (ERIC). ECRIN provides support for the development and implementation of multinational clinical research projects in Europe. It currently has seven Member Countries and two Observer Countries. Some of ECRIN's activities are related to RD, especially clinical trials in RD and Paediatrics (PedCRIN) for development of new therapies in RD. ECRIN has built partnership with the RD community including Orphanet. Patients' organisations are fully integrated into their projects as they are involved in the evaluation of the projects in general or in specific clinical trials upon request of the coordinator or Principal Investigators. In addition, the involvement of patients' organisations is part of the ECRIN evaluation criteria to support project and it is recommended to use patients' organisations when available. However, no concrete plans for collaboration with ERNs have been made yet.

PEDCRIN is a recent (2017) project bringing together ECRIN and the founding partners of the European Paediatric Clinical Trial Research Infrastructure (EPCT-RI) to develop the necessary tools and capacity to enhance the quality, safety, efficacy and ethical standards of multinational, non-commercial paediatric clinical trials. The PEDCRIN project is not yet mature and not RD-specific but the tools, methods, and services that

will be developed for academic and independent investigators for conducting multinational neonatal and paediatric clinical trials could also be adopted for RD.

INFRAFRONTIER is the European research infrastructure for the development, phenotyping, archiving and distribution of model mammalian genomes. Formed by more than 25 research centres in 14 European countries and one in Canada, it offers open access to unique scientific platforms, resources and services. INFRAFRONTIER uses ORDO but is not fully FAIR yet as it does not have APIs for the moment.

Data sharing:

RD-Nexus is an array of networks making their assets safely discoverable, without ever exposing the data content of the asset itself. Control over the data (queries, responses, etc.) remain in the hands of the data owner. RD-Nexus is an RD-specific that is still being developed.

Omics:

neXtProt is a comprehensive human-centric discovery platform offering its users a seamless integration of and navigation through protein-related data. neXtProt complements UniProtKB/Swiss-Prot with genomics, transcriptomics and proteomics data from various sources. It has a web interface, dedicated tools for the proteomics community, a SPARQL endpoint allowing advanced data querying, and an API allowing easy connection with other life science resources and tools. However, neXtProt uses neither ORDO nor a standard linking with ORDO, which prevents it from being completely RD-optimized.

The PRoteomics IDentifications (PRIDE) database is one of the founding members of the global ProteomeXchange (PX) consortium. In addition to be an archive, PRIDE is starting to systematically re-analyse public proteomics datasets so that the data is disseminated to added-value resources including Ensembl, UniProt and Expression Atlas. However, PRIDE does not use ORDO.

The UniProt Knowledgebase is a collection of sequences and annotations for currently 160 million proteins across all branches of life. As a project started in the mid 80's, it is a well-developed and mature project with cross-referenced annotations from several open-source sources including Orphanet. It is a sustainable and FAIR resource with persistent IDs, APIs, and licencing issues long solved.

- Less than 20% of core criteria fulfilled:

Resources: Euro-Biolmaging, EATRIS-ERIC, IDEAL, I-Stem, Cellosaurus, ChemBioNet, and EU-OPENSREEN.

Catalogue:

EuroBioImaging (EuBI) is a recent (2018) European-wide research infrastructure of 29 imaging facilities across 10 European countries. EuBI offers access to cutting-edge imaging technologies, expert training and support, and image processing and storage. EuBI does not conduct activities specifically related to RD and no RD-specific tools are planned to be developed in the future. However, their current tools can be adapted to RD research. It is also not involved with any RD patients' organization.

Clinical and translational research:

EATRIS is a resource making translation of scientific discoveries into medical products more effective to improve human health and quality of life. EATRIS is a well-embedded resource within the RD research landscape, has a RD working group, and provides support to researchers developing new RD therapies via four current projects in RD benefitting from EATRIS services.

Projects:

The IDEAL project's aim was to refine the statistical methodology in small population group trials. All its activities were intended for RD, but the project is now over, results have been published, and the research is discontinued.

Solve-RD is a recent (2018) project whose ambitions are to solve a large number of RD without known molecular cause as well as improve diagnostics of RD through the implementation of a "genetic knowledge web". Solve-RD's tool developments contribute to diagnostics in RD beyond the exome and will provide bioinformatic tools needed to solve ca. 50% of unsolved patients in RD. Solve-RD uses HPO terms, FAIR principles for the data, and is fully interconnected with RD-Connect PhenoTips where all clinical data are collected. The project is well optimized for RD research.

Models:

I-stem is a research laboratory part of the French Biotherapies Institute specialised in RD and funded directly by the French Muscular Dystrophy Association (AFM). I-stem is dedicated to the development of treatments based on the potentials offered by pluripotent stem cells and applicable to rare diseases of genetic origin. Although completely RD-oriented, I-stem does not possess an openly-accessible platform with specific standards for data sharing, is not interoperable with other resources and thus is not a FAIR resource. I-stem does not have plans for future collaborations with ERNs.

Cellosaurus is a knowledge resource on cell lines. With now more than 115,000 cell lines from vertebrates (with an emphasis on human, mouse and rat) and invertebrates-it is the most comprehensive cell line resource. It is cross-referenced to about 90 resources

and can be downloaded in a variety of formats. However, Cellosaurus does not use ORDO as a standard and has no API to make it a fully FAIR resource, it is therefore not fully RD-optimized.

Pharmacology/Molecular screening:

ChemBionet is an association created in 2004 whose purpose is the sharing of compounds and screening access across participating laboratories. It does not conduct activities specifically related to RD, but its generic tools can be adapted for RD. No RD-specific tools and/or methods are currently planned to be developed by ChemBionet although such developments could be envisioned if requested. Also, the sustainability of ChemBionet is still being studied-the resource might become the German branch of EU-OPENSREEN.

EU-OPENSREEN is a recently-created (2018) European research infrastructure consortium and is still being developed. Its primary objective is to support scientists in identifying, developing and providing novel chemical compounds that play an important role in validating novel drug targets during early drug discovery. The project is not yet mature.

Particular resources/Projects:

GO-FAIR, FAIR sharing, FAIR evaluator, EUPID, and PPRL. These tools/platforms are meant to support, link, or provide a service to research resources or to provide standards that can be used by researchers and thus cannot be classified in the same way as other resources. However, EUPID has been promoted by the GA4GH-IRDiRC PPRL task force¹.

Resources that did not respond:

The following resources were contacted but did not answer the surveys: ADA-M, Beacon Project - GA4GH, CIDSTEM, CORBEL, COSMOS (EBI), EGA, ELIXIR, EOSCpilot, EUPID, EuRenOmics, GRIP, HGNC, HPO, INFRAFRONTIER, INSPIRE, LOVD, NeurOmics, Orphamizer, PhenoMeNal, TREAT-NMD Patient Registries, and WormBase. For a description of these resources, please refer to ANNEX I.

¹ Baker, Dixie, et al. "Privacy-preserving linkage of genomic and clinical data sets." *IEEE/ACM Transactions on Computational Biology and Bioinformatics* (2018).

Table 1. Evaluation of resources based on the 2017 and 2019 surveys on resources.

The table allows to evaluate their "readiness" for RD research. Green "+" = RD-optimized; Yellow "±" = Criterion soon to be fulfilled; Orange "-" = Not yet optimized; Grey square = information not available; Blue square: non-applicable. **Red block** = core criteria; **Orange block** = Additional criteria (see text for details).

		RD-Standards	Access model	FAIR resource/API	Interoperability	RD-specific	Ext. quality assess.	FAIR-sharing-reco.	Sustainability	Collaborations	Maturity	Transparency	Future RD involv.	Involvement of PO
Catalogues	BBMRI-ERIC	+	+	+	+	-	-	-	+	+	+	+	+	+
	EuroBioBank	+	+	+	+	+	-	-	±	+	±	+	+	+
	EU RD Platform	+	+	+	+	+	-	-	+	+	±	+	+	-
	Orphanet	+	+	+	+	+	+	-	±	+	+	+	+	+
	RaDiCo	+	+	-	+	+	-	-	±	+	+	+	+	+
	RD-Connect	+	+	+	+	+	-	-	±	+	±	-	+	+
	RD-Connect Samp. Cat	+	+	+	+	+	-	-	±	±	±	±	±	-
	Euro-Biolmaging				+	-	-	-	±	-	-	-	±	-
Translational/ clinical research	EORTC-SPECTA				+	+	-	-	+	+	+	+	+	+
	ECRIN				+	+	-	-	+	-	±	+	+	+
	PEDCRIN				+	+	-	-	+	-	±	+	+	+
	INFRAFRONTIER	+	+	-		-	-	-		+				
	EATRIS-ERIC				-	+	-	-	+	+	±	+	+	+
Projects/Task Forces/Tools	FAIR evaluator					-	-	-		+	±	+	+	-
	FAIRsharing					-	-	-	±	-	±	+	±	-
	GO-FAIR					-	-	-	±	+	±	+	+	+
	EUPID					-	-	-		+	±	+	+	-
	IDEAL					+	-	-	-	+	+	+	+	+
	PPRL Task Force					+	-	-	±	+	-	-	-	+
	SOLVE-RD					+	-	-	-	+	-	-	+	+
Onto- logies	HPO	+	+	+	+	-	+	+		+				
	ORDO	+	+	+	+	+	+	-	±	+	+	+	+	+
Data sharing	DECIPHER	-	+	+	+	+	+	-	+	+	+	+	+	-
	RD-Connect GPAP	+	+	+	+	+	+	-	±	+		+	+	
	RD-NEXUS	+	+	-		+	-	-		+				
Omics	Ensembl	+	+	+	+	-	+	+		+				
	MetaboLights	+	+	+	+	-	-	-		+				
	neXtProt	-	+	+	+	-	-	-		+				
	PRIDE	-	+	+	+	-	+	+		+				
	UniProt	-	+	+	+	-	+	+	+		+	+	+	-
Models	I-Stem				-	+	-	-	+	-	+	+	+	+
	Cellosaurus	-	+	-		-	-	-		+				
	hPSCreg	+	+	+	+	-	-	-		+				
Phar- maco.	ChemBioNet				+	-	-	-	±	-	+	+	±	-
	EU-OPENSREEN					-	-	-	+	-	-	+	±	-

5 Conclusions

The categorization system used above allows to see quickly the level of optimization of RD resources surveyed. It is important to note that all resources, including the group composed of resources fulfilling 80-100% of core criteria, could currently increase this level optimization for RD research—no resource has fulfilled all criteria of the survey. It is also worth mentioning that it is difficult to characterize all resources equally considering their variety of use and purpose. The classification scheme (core vs additional) applied to the evaluation criteria could be debated. Regardless of these points of discussion, this analysis provides a clearer picture of the work that remains to be done on each resource to increase their usability in the RD context and what the EJP RD can contribute to each one of them. However, it is worth to remind that this categorization has been possible only for 35 responding resources.

One look at the current RD resources landscape shows that it has evolved and diversified enough to make one feel that the time is ripe to bring all these resources together to synergise their actions—this observation crosses borders as attested by initiatives such as the recent national platform for Rare Diseases Data Registry of Japan¹ (RADDAR-J) or the Open Platform for Rare Diseases² (OPFORD) in India. In that context, the EJP RD's upcoming virtual platform is particularly timely. In order to improve the usefulness and usability of these resources, four main axes should be focused on in order to ensure that the minimum requirement in RD research are met. These four axes are (1) data visibility and discoverability, (2) data sharing/access, (3) data interoperability and alignment (RD-specific terminology), and (4) data sustainability.

Ensuring interoperability between resources is key to improving research efficiency and the lack of compatibility between tools has been suggested to be a major concern in users³. In that respect, compliance with FAIR principles is promoted within EJPRD. In the RD field, to achieve semantic interoperability, it seems necessary to ensure that resources involved in RD use ORDO as their standard for RD nomenclature or ensure full interoperability with it⁴. The Orphanet RD nomenclature is the only RD-specific nomenclature and has been recommended by the European Union Committee of Experts on Rare Diseases (EUCERD) as a coding system within the Member States' health information systems and is also described as an important

¹ Yamaguchi, I., Furusawa, Y., Kawaguchi, T., Yagishita, N., Tanzawa, K., Yamano, Y., & Matsuda, F. (2019). Establishment of a Comprehensive Information Infrastructure and a Support Organization for Rare Disease Research in Japan (RADDAR-J). *Studies in Health Technology and Informatics*, 264, 1080–1083. <http://doi.org/10.3233/SHTI190391>

² Open Platform for Rare Diseases (OPFORD) website: <https://opford.org>

³ Coward, E. (2017). Rare diseases data resources and tools survey: Summary of results [Elixir report]. <https://drive.google.com/file/d/0B6ue78XJUF74b0dITjZfQ2EtaVU/view>

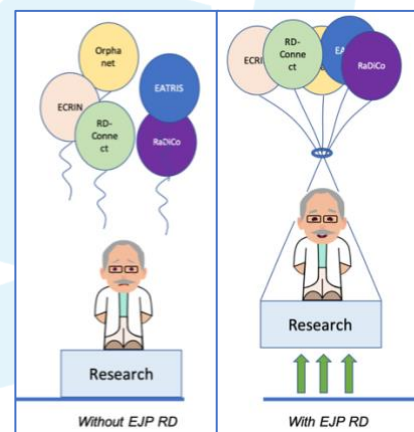
⁴ Montserrat Moliner, A., & Waligóra, J. (2013). The European union policy in the field of rare diseases. *Public Health Genomics*, 16(6), 268–277. <http://doi.org/10.1159/000355930>

starting point when investigating RD_{1,2}. It is therefore a common language allowing to bridge health and research.

The EJP RD will allow an increased networking among resources and partners, an allocation of budget for necessary standardisation to make them fully RD research-compatible, and a better integration into the whole RD landscape. Thanks to the EJP RD, expertise in FAIRification, data standards, and RD knowledge are concentrated in space and time, which provides the opportunity for a cooperative improvement and evolution of platforms and tools towards a more fluid experience in sharing and consulting data as well as an efficient collaboration among RD partners. As a major project and one of the world's leading genomic data initiatives, the EJP RD has recently joined the Global Alliance for Genomics and Health (GA4GH) as a 2019 Driver Project. As such, the EJP RD will work collaboratively to develop and pilot GA4GH standards for sharing genomic and health related data. This collaboration will contribute to data harmonization as well as data and knowledge sharing across national borders. It will pave the road for the building of global-scale frameworks and will broaden the network of RD research initiatives that will provide guidance to the RD community for years to come.

All the EJP RD participants will benefit from the concerted and focused effort to bring together all relevant RD resources and make them accessible via a single, unified virtual platform. While currently, researchers must grasp a variety of dispersed resources making them waste precious time in the form of resources search and access, lack of tools interoperability, data licensing issues, and data discrepancies, research in RD will be uplifted thanks to the EJP RD and the virtual platform which will allow pooling resources in a single place as well as ensure their unicity, ease of access, and an overall better inter-connectivity among them.

Figure 3. Expected effect of EJP RD on research efforts.

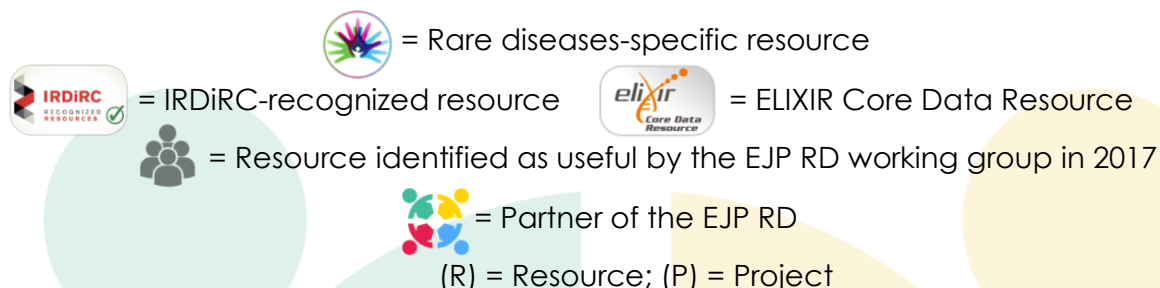


¹ Pogue, R. E., Cavalcanti, D. P., Shanker, S., Andrade, R. V., Aguiar, L. R., de Carvalho, J. L., & Costa, F. F. (2018). Rare genetic diseases: update on diagnosis, treatment and online resources. *Drug Discovery Today*, 23(1), 187–195. <http://doi.org/10.1016/j.drudis.2017.11.002>

² Vandeborne, L., van Overbeeke, E., Doms, M., De Beleyr, B., & Huys, I. (2019). Information needs of physicians regarding the diagnosis of rare diseases: a questionnaire-based study in Belgium. *Orphanet Journal of Rare Diseases*, 14(1), 99. <http://doi.org/10.1186/s13023-019-1075-8>

ANNEX I. List of research resources for rare diseases

106 resources have been catalogued based on the combined work of the 2017 Working Group (WG) and the Orphanet Report Series on RD research infrastructures¹. EJP RD partner-resources and WG-identified resources are indicated with a specific symbol (see below).



ADA-M (Automatable Discovery and Access Matrix) (R)



Comprehensive information model that provides the basis for producing structured metadata "Profiles" of regulatory conditions, thereby enabling efficient application of those conditions across regulatory spheres. ADA-M was created to help increase the efficiency of resource discovery and access, by promoting responsible recording, versioning, communication, querying, and actioning of resource sharing plans.

<http://www.irdirc.org/responsible-sharing-of-biomedical-data-and-biospecimens-via-the-automatable-discovery-and-access-matrix-ada-m/>

Data access/Sharing; Coverage: Global

ADOPT-BBMRI (P)



The ADOPT BBMRI-ERIC proposal aims at boosting and accelerating implementation of BBMRI-ERIC and its services. Its main deliverables are designed to complete or launch the construction of key Common Services of the Research Infrastructure as required for ESFRI-projects "under implementation", reflecting the targets of the European Research Area (ERA). The project ended in March 2019.

<http://www.bbmri-eric.eu/scientific-collaboration/adopt-bbmri-eric/>

Catalogue [biobanks]; Coverage: Europe

¹ Orphanet. 2019. "List of Research Infrastructures useful to Rare Diseases in Europe":
https://www.orpha.net/orphacom/cahiers/docs/GB/Research_Infrastructures_for_rare_diseases_in_Europe.pdf

APTEEUS (R) APTEEUS



Clinical stage biotech company addressing monogenetic disorders and focused on the molecular causes of diseases, the discovery, and the development of Disease Modifying Therapies.

<https://apteus.fr/en/>

Translational research; Coverage: National (FR)

Atlantic Gene Therapies [Biotherapies Institute] (R)



See *Biotherapies Institute* entry.

BBMRI-ERIC (Biobanking and BioMolecular resources Research Infrastructure - European Research Infrastructure Consortium) (R)



European research infrastructure for biobanking. BBRMI-ERIC brings together all the main players from the biobanking field – researchers, biobankers, industry, and patients – to boost biomedical research. BBRMI-ERIC offers quality management services, support with ethical, legal and societal issues, and a number of online tools and software solutions.

<http://www.bbmri-eric.eu>

Catalogue [biobanks] ; Coverage: Europe

Beacon Project - GA4GH (P) Beacon Network



Genetic mutation sharing platform developed by the Global Alliance for Genomics and Health. The Beacon Network is a search engine across the world's public beacons (institutions' web services used to share genetic data). It enables global discovery of genetic mutations, federated across a large and growing network of shared genetic datasets.

<https://beacon-network.org/#/>

Data access/Sharing; Coverage: Global

BioMedBridges [See also CORBEL] (P)



Joint effort of twelve biomedical sciences research infrastructures to develop the shared e-infrastructure—the technical bridges—to allow data integration in the biological, medical, translational and clinical domains and thus strengthen biomedical resources in Europe. The project successfully concluded in December 2015, delivering the basis and impetus for its follow-on sister project CORBEL.

www.biomedbridges.eu/

Data access/Sharing; Coverage: Global

Biotherapies Institute (R)


Union of three laboratories (Genethon, Institute of Myology, and I-Stem), their knowledge and expertise at the service of projects carried out by the laboratories of the Biotherapies Institute and/or teams of excellence supported by the AFM-Telethon. This expertise ensures the implementation and effective coordination of discovery projects and development of new drugs with the aim of starting a clinical trial.

www.institut-biotherapies.fr

Translational research; Coverage: National (FR)

C4C - Conect4Children (P)


Large collaborative European network that aims to facilitate the development of new drugs and other therapies for the entire paediatric population.

<https://conect4children.org>

Clinical trials; Coverage: Europe

Cafe Variome (R)


Flexible web-based, data discovery tool that can be quickly installed by any biomedical data owner to enable the “existence” rather than the “substance” of the data to be discovered. Cafe Variome has been designed for use with all sensitive biomedical data, whether this be genomic variants or cohort data.

<https://www.cafevariome.org>

Data access/Sharing; Coverage: Global

Cellosaurus (R)


Knowledge resource on cell lines. It attempts to describe all cell lines used in biomedical research.

<https://web.expasy.org/cellosaurus/>

Models (animal, cells); Coverage: Global

ChemBioNet (R)


Resource network providing chemists with bioprofiles for their unique synthetic molecules and biologists developing unique assay systems, with access to high throughput technologies to identify compounds useful for dosage dependent, temporally or locally controlled interference with biological functions.

www.chembionet.info

Pharmacology/Molecular screening; Coverage: National (DE)

ChEMBL (Network for Academic Chemical Biology Research) (R)


Manually curated database of bioactive molecules with drug-like properties. It brings together chemical, bioactivity and genomic data to aid the translation of genomic information into effective new drugs.

<https://www.ebi.ac.uk/chembl/>

Pharmacology/Molecular screening; Coverage: Europe

CIDSTEM (Centro Interdipartimentale Cellule Staminali e Medicina Rigenerativa [Interdepartmental Centre Stem Cells and Regenerative Medicine]) (R)



High Technology Network ensuring the conversion of knowledge generated by regional public research into technologies and processes, then into marketable products and services for businesses.

http://www.cidstem.unimore.it/chi_siamo.html

Models (animal, cells); Coverage: National (IT)

ClinGen (Clinical Genome Resource) (R)



National Institutes of Health (NIH)-funded resource dedicated to building a central resource that defines the clinical relevance of genes and variants for use in precision medicine and research.

<https://www.clinicalgenome.org/>

Mutations database; Coverage: Global

Cliniface (former 3D-FAST) (R)



Tool that improves clinical workflows and support the diagnostic decision-making process in rare diseases by performing fast and accurate analysis of input 3D scans of patient faces stored in a variety of formats. The Cliniface platform is a desktop software application for Windows and Linux providing a suite of tools and visualisation technologies to support the analysis and processing of 3D facial images. The application is aimed primarily at medical professionals who seek to leverage 3D facial images to explore the richness of the facial surface to visualise, identify, and communicate cues of syndromic or surgical relevance.

<http://www.crcsi.com.au/assets/Program-4/Health/Posters-2016/P4.406-CliniFace-Project-2016.pdf>

<https://cliniface.org>

Phenotype/Clinical features; Coverage: Global

ClinVar (R)

ClinVar
Clinically relevant variation



National Institutes of Health (NIH)-funded freely accessible, public archive of reports of the relationships among human variations and phenotypes, with supporting

evidence.

<https://www.ncbi.nlm.nih.gov/clinvar/>

Mutations database; Coverage: Global

COMET Initiative (Core Outcome Measures in Effectiveness Trials Initiative) (P)



COMET brings together people interested in the development and application of agreed standardised sets of outcomes, known as 'core outcome sets' (COS). These sets represent the minimum that should be measured and reported in all clinical trials of a specific condition, and are also suitable for use in clinical audit or research other than randomised trials.

Clinical trials; Coverage: Europe

CORBEL [Previously BioMedBridge] (Coordinated Research Infrastructures Building Enduring Life-science Services) (P)



Platform for harmonised user access to biological and medical technologies, biological samples and data services required by cutting-edge biomedical research.

<http://www.corbel-project.eu/home.html>

Data access/Sharing; Coverage: Europe

COSMOS (EBI) (COordination Of Standards In MetabOmicS) (P)



Consortium of leading European groups in Metabolomics interfacing with all interested players in Metabolomics world-wide in the Metabolomics community and beyond.

<http://cosmos-fp7.eu/index.html>

Metabolomics; Coverage: Europe

CRIGH (The Clinical Research Initiative for Global Health) (P)



Consortium of research institutions and organisations supporting international collaboration on clinical research.

<http://www.ecriin.org/projects/crigh>

Clinical trials; Coverage: Global

CTSR (Care and Trial Site Registry) (R)



CTSR helps the pharmaceutical industry and clinical investigators select trial sites as well as to help to identify potential partners for upcoming research projects in neuromuscular and neurodegenerative diseases. CTSR provides information

relevant to clinical studies (such as personnel, facilities and patient population) and to the assessment as centres of expertise according to the EUCERD criteria.

<https://ctsr.uniklinik-freiburg.de/ctsr/index>

Clinical trials; Coverage: Europe

DataSharing (P)



Pointer to currently available sources of information and software, for analysis, interpretation and sharing of genetic data.

https://datasharing-101.le.ac.uk/DataSharing_101/page0/page0.php

Data access/Sharing; Coverage: Europe

DECIPHER (Database of genomic variation and Phenotype in Humans using Ensembl Resources) (R)



Interactive web-based database which incorporates a suite of tools designed to aid the interpretation of genomic variants.

<https://decipher.sanger.ac.uk/>

Phenotype/Clinical features; Coverage: Global

EATRIS (European Advanced Translational Research InfraStructure in medicine) (R)



European Research Infrastructure Consortium supporting translational research by providing guidance in the steps to be taken, by means of clinical, biological and technological expertise available within the infrastructure.

<http://eatris.eu/>

Translational research; Coverage: Europe

ECARUCA (European Cytogeneticists Association Register of Unbalanced Chromosome Aberrations) (R)



ECARUCA collects and provides detailed, curated clinical and molecular information on rare unbalanced chromosome aberrations.

<https://omictools.com/ecaruca-tool>

Genomics; Coverage: Europe

ECRIN (European Clinical Research Infrastructure Network) (R)



ECRIN is a European research Infrastructure consortium (ERIC) that links scientific partners and networks across Europe to facilitate multinational clinical research. It provides sponsors and investigators with advice, management services and tools to overcome hurdles to multinational trials and enhance collaboration.

<https://www.ecri.org>

Clinical trials; Coverage: Europe

EGA (European Genome-phenome Archive) (R)



EGA provides a service for the permanent archiving and distribution of personally identifiable genetic and phenotypic data resulting from biomedical research projects.

<https://ega-archive.org/>

Genomics; Coverage: Europe

ELIXIR (R)



Intergovernmental organisation that brings together and coordinates life science resources, including databases, software tools, training materials, cloud storage and supercomputers, from across Europe so that they form a single infrastructure.

<https://www.elixir-europe.org/>

Data access/Sharing; Coverage: Europe

ELIXIR - AAI (ELIXIR AAI: Authentication and Authorisation Infrastructure) (R)



ELIXIR - AAI enables researchers to use their home organisation credentials or community or commercial identities (e.g. ORCID, LinkedIn) to sign in and access data and services they need. It also allows service providers (both in academia and industry) to control and manage access rights of their users and create different access levels for research groups or international projects.

<https://elixir-europe.org/services/compute/aaai>

Data access/Sharing; Coverage: Europe

ELIXIR Biotools (R)



Discovery portal for bioinformatics software information.

<https://elixir-europe.org/services>

Data capture/analysis; Coverage: Europe

EMBL-EBI (European Molecular Biology Laboratory - European Bioinformatics Institute) (R)



EMBL-EBI makes the world's public biological data freely available to the scientific community via a range of services and tools, performs basic research and provides professional training in bioinformatics. Maintains the world's most comprehensive range of freely available and up-to-date molecular data resources to share data, perform complex queries and analyse the results in different ways.

<https://www.ebi.ac.uk>

Data capture/analysis; Coverage: Europe

Ensembl (EBI) (R)



Genome browser for vertebrate genomes that supports research in comparative genomics, evolution, sequence variation and transcriptional regulation.

<https://www.ensembl.org/index.html>

Genomics; Coverage: Global

EORTC (European Organisation for Research and Treatment of Cancer) (R)



The future of cancer therapy



Non-profit cancer research organisation that coordinates and conducts international translational and clinical research to improve the standard of cancer treatment for patients.

www.eortc.org

Clinical trials; Coverage: Europe

EOSCpilot (European Open Science Cloud for Research Pilot Project) (P)



The European Open Science Cloud for Research Pilot Project



EOSCpilot aims to offer a virtual environment with open and seamless services for storage, management, analysis and re-use of research data, across borders and scientific disciplines by federating existing scientific data infrastructures, today scattered across disciplines and Member States.

<https://eoscipilot.eu/>

Data access/Sharing; Coverage: Europe

ERNs (European Reference Networks) (R)



Virtual networks involving healthcare providers across Europe. They aim to facilitate discussion on complex or rare diseases and conditions that require highly specialised treatment, and concentrated knowledge and resources, as well as to disseminate best practices across Europe for the treatment of RD.

https://ec.europa.eu/health/ern/networks_en

RD-integrated resources; Coverage: Europe

ESGI (European Sequencing and Genotyping Infrastructure) (R)



Network of leading European genome research centres that aims to integrate the capacity for large-scale genomic analysis and technical expertise in sequencing and genotyping technology among the partner institutes.

<https://www.esgi-infrastructure.eu/>

Genomics; Coverage: Europe

eTRIKS (ELIXIR) (R)



Repository that centralizes ongoing and past Innovative Medicines Initiatives (IMI) project level metadata. eTRIKS/ELIXIR-LU Data Catalogue is a metadata repository linking the massive data available in a global system that can be optimally leveraged to improve biomedical research. This is a collaborative project between eTRIKS and ELIXIR-Luxemburg Node.

<https://www.etricks.org>

Data access/Sharing; Coverage: Europe

EU-OPENSREEN (European Infrastructure of Open Screening Platforms for Chemical Biology) (R)



Open access to a uniquely broad range of high technologies and tools for the systematic screening of chemical substances for their biological effects.

<https://www.eu-openscreen.eu>

Pharmacology/Molecular screening

EUPID (European Patient IDentity management) (R)



EUPID facilitates secondary use of datasets in Biomedical Research and Healthcare by (1) preventing duplicate registration of patients, (2) providing distinct pseudonyms for patients in different contexts, (3) preserving the possibility for re-identification by a trusted third party, and (4) keeping a protected link between the different pseudonyms in the background.

[Linked to RD-Connect/GA4GH Privacy-Preserving Record Linkage (PPRL) project]

<https://eupid.eu>

Data access/Sharing; Coverage: Europe

EuRenOmics (European Consortium for High-Throughput Research in Rare Kidney Diseases) (P)



Central 'renal phenome' database integrating clinical information from local and regional patient registries as well as a unique kidney-focused bioinformatic analysis pipeline; Prognostic biomarkers

<https://eurenomics.eu>

Biomarkers; Coverage: Europe

EuroBioBank (R)



Network of biobanks that stores and distributes quality DNA, cell and tissue samples for scientists conducting research on rare diseases. The network consists of 25 RD biobanks located in 11 countries. It is the biobank network of RD-Connect.

<http://www.eurobiobank.org>

Catalogue [biobanks] ; Coverage: Europe

EuroBioImaging (R)



EuroBioImaging provides open physical user access to a broad range of state-of-the-art technologies in biological and biomedical imaging for life scientists. EuroBioImaging will also offer image data support and training for infrastructure users and providers.

<http://www.eurobioimaging.eu/>

Imaging; Coverage: Europe

EuroStemCell (European Consortium for Communicating Stem Cell Research) (R)



EuroStemCell provides independent, expert-reviewed information and road-tested educational resources on stem cells and their impact on society.

<http://www.eurostemcell.org/>

Models (animal, cells) ; Coverage: Europe

Exomiser (R)



Java program that finds potential disease-causing variants from whole-exome or whole-genome sequencing data.

<https://github.com/exomiser/Exomiser>

Data capture/analysis; Coverage: Europe

Face2Gene (R)



Suite of phenotyping applications that facilitate comprehensive and precise genetic evaluations.

<https://www.face2gene.com>

Data capture/analysis; Coverage: Global

FAIR data services (GO FAIR) (GO FAIR - The internet of FAIR Data & Services) (R)



GO FAIR supports stakeholders in applying FAIR data principles for their data (Findable, Accessible, Interoperable, Reusable for humans and computers). Services encompass (i) software to aid FAIR data management; (ii) implementation services that provide data stewardship expertise to novices (includes on-site support, training, and liaising between users and software engineers), (iii) organizational services to support the organization of FAIR projects (e.g. via FAIR data project blueprints for planning and budgeting FAIRification), 'Bring Your Own Data' workshops, and the organization of Global Open FAIR networks.

<https://www.go-fair.org>

Data access/Sharing; Coverage: Global

**FAIRsharing (ELIXIR)
(BioSharing) (R)**



A curated, informative and educational resource on data and metadata standards, inter-related to databases and data policies. FAIRsharing guides consumers to discover, select and use these resources with confidence, and producers to make their resource more discoverable, more widely adopted and cited.

<https://fairsharing.org>

Data access/Sharing; Coverage: Global

France Genomique (R)



France Genomique offers the public and private scientific community the highest level of expertise and skills, as well as project support.

<https://www.france-genomique.org/?lang=en>

Genomics; Coverage: National (FR)

**GeneOntology (Gene Ontology
knowledgebase) (R)**



World's largest source of information on the functions of genes. This knowledge is both human-readable and machine-readable and is a foundation for computational analysis of large-scale molecular biology and genetics experiments in biomedical research.

<http://www.geneontology.org/>

Ontology; Coverage: Global

Genethon (R)



Design and development of gene therapy treatments for rare diseases. Genethon is developing therapies for rare neuromuscular diseases, immune system or blood disorders, eye disorders and liver diseases.

www.genethon.fr

Translational research; Coverage: National (FR)

Genomiser (R)



Phenotype-based tool that is able to score the relevance of variation in the non-coding genome and also to associate regulatory variants to specific Mendelian diseases.

<https://hpo.jax.org/app/tools/genomiser>

Data capture/analysis; Coverage: Global

**GRiP (Global Research in
Paediatrics) (P)**



GRiP is a network among the main international paediatric centres – with over 1,000 structures involved in Europe, the U.S., Canada, Japan and other countries – in the form of a “Network of Excellence” meant to stimulate and facilitate the development and safe use of medicines for children.

<http://www.grip-network.org>

Clinical trials; Coverage: Global

HGNC (HUGO [Human Genome Organisation] Gene Nomenclature Committee) (R)



The HGNC is responsible for approving unique symbols and names for human loci, including protein coding genes, ncRNA genes and pseudogenes, to allow unambiguous scientific communication.

<https://www.genenames.org>

Ontologies/Standards; Coverage: Global

HPO (Human phenotype ontology) (R)



HPO provides a standardized vocabulary of phenotypic abnormalities encountered in human disease. Each term in the HPO describes a phenotypic abnormality. HPO is currently being developed using the medical literature, Orphanet, DECIPHER, and OMIM.

<https://hpo.jax.org/app/>

Ontology; Coverage: Global

hPSCreg (Human Pluripotent Stem Cell Registry) (R)



Freely accessible global registry for human pluripotent stem cell lines (hPSC lines). The registry allows searching for cell lines and for information available about these cell lines. New cell lines can be registered and information to already registered cell lines can be added.

<https://hpscereg.eu>

Models (animal, cells) ; Coverage: Global

I-Stem (Institute for Stem cell Therapy and Exploration of Monogenic diseases) (R)



Largest French laboratory for research and development dedicated to human pluripotent stem cells, of embryonic origin or obtained by reprogramming gene. I-Stem is part of the Biotherapy Institute for Rare Diseases, which includes so far the three centres of research and development funded directly by the AFM Telethon.

www.istem.eu

Models (animal, cells) ; Coverage: National (FR)

i2b2 (Informatics for Integrating Biology & the Bedside) (R)

The i2b2 tranSMART Foundation is a non-profit foundation developing an open-source / open-data community around the i2b2, tranSMART and OpenBEL translational research platforms. The i2b2 tranSMART Foundation enables effective collaboration for precision medicine, through the sharing, integration, standardization and analysis of heterogenous data from healthcare and research; through engagement and mobilization of a life sciences focused open-source, open-data community.

<https://www.i2b2.org>

Data access/sharing; Coverage: Global

IDEAL (Integrated DEsign and Analysis of small population group trials) (P)

Innovative methodologies improving clinical trials in the setting of small sample population groups. Focussed on assessment of randomization procedures, extrapolating dose-response information, investigation of adaptive designs, optimal designs in mixed models, pharmacogenetic designs, simulation of clinical trials, genetic factors influencing the response, decision analysis and biomarker surrogate endpoints.

<http://www.ideal.rwth-aachen.de/>

Clinical trials; Coverage: Europe

IMPC (International Mouse Phenotype Consortium) (R)

International effort to identify the function of every protein-coding gene in the mouse genome. IMPC creates a comprehensive catalogue of mammalian gene function that is freely available for researchers.

<https://www.mousephenotype.org>

Models (animal); Coverage: Global

INCF (International Neuroinformatics Coordinating Facility) (R)

INCF promotes the field of neuroinformatics and aims to advance data reuse and reproducibility in global brain research. INCF is an independent international facilitator catalysing and coordinating the global development of neuroinformatics and advancing training in the field. INCF participates in a number of international collaborations in relation to neuroinformatics training, coordination of large-scale projects, scientific publishing, and outreach.

<https://www.incf.org>

Data capture/analysis; Coverage: Global



INFRAFRONTIER (R)



Archiving and Distribution of Mouse Models; Systemic Phenotyping of Mouse Models (Transnational Access service); Mouse Production from ES Cells (Transnational Access service); Training and Consulting.

<https://www.infrafrontier.eu>

Models (animal, cells) ; Coverage: Europe

INSPIRE (Innovation in Small Populations Research) (P)



Clinical trial design and analysis methodology in rare diseases or small populations defined, for example, by a rare genetic marker.

<https://warwick.ac.uk/fac/sci/med/research/hscience/stats/completedprojects/inspire/>

Clinical trials; Coverage: Europe

Instruct-ERIC (Instruct-European Research Infrastructure Consortium) (R)



Pan-European distributed research infrastructure making high-end technologies and methods in structural biology available to users. Instruct-ERIC provides open access to cutting edge structural biology, specifically supporting research that uses integrated approaches and technologies.

<https://instruct-eric.eu>

Data capture/analysis; Coverage: Global

ISPOR (International Society for Pharmacoeconomics and Outcomes Research) (R)



Leading professional society for health economics and outcomes research (HEOR). The Society's mission is to promote HEOR excellence to improve decision making for health globally.

<https://www.ispor.org/home>

Translational research; Coverage: Global

IUPHAR (International Union of Basic and Clinical Pharmacology) (R)



IUPHAR (1) provides guidelines for the nomenclature and classification of all the (human) biological targets, including all the targets of current and future prescription medicines; (2) facilitates the interface between the discovery of new sequences from the Human Genome Project and the designation of the derived entities as functional biological targets and potential drug targets; (3) designates polymorphisms and variants which are functionally important; (4) develops an authoritative and freely available, global online resource, the IUPHAR database (<https://www.iuphar-db.org>).

<http://www.guidetopharmacology.org/>

Pharmacology/Molecular screening; Coverage: Global

JRC RD EU Platform (Joint Research Centre European Platform on Rare Disease Registration) (R)



The EU RD Platform copes with the enormous fragmentation of rare disease (RD) patients' data contained in hundreds of registries across Europe. The Platform makes RD registries' data searchable and findable, thus increasing visibility for each registry, maximising the value of each registry's information and enabling extended use and re-use of registries' data. This is ensured by the European RD Registry Infrastructure (ERDRI), which supports existing registries and the creation of new registries. In addition to ERDRI, the EU RD Platform includes a data repository composed of the European RD Registry Data Warehouse (under preparation), the JRC-EUROCAT Central Registry and the JRC-SCPE Central Registry.

<https://eu-rd-platform.jrc.ec.europa.eu>

Catalogue [registries/biobanks]; Coverage: Europe

LOVD (Leiden Open Variation Database) (R)



LOVD provides a flexible, freely available tool for Gene-centered collection and display of DNA variants. LOVD 3.0 extends this idea to also provide patient-centered data storage and storage of NGS data, even of variants outside of genes.

www.lovd.nl

Mutations database; Coverage: Global

MARRVEL (Model organism Aggregated Resources for Rare Variant ExpLoration) (R)



MARRVEL facilitates the use of public genetic resources to prioritize rare human gene variants for study in model organisms. To facilitate the search process and gather all the data in a simple display, data is extracted from human data bases (OMIM, ExAC, ClinVar, Geno2MP, DGV, and DECIPHER) for efficient variant prioritization. The protein sequences for eight organisms (*S. cerevisiae*, *S. pombe*, *C. elegans*, *D. melanogaster*, *D. rerio*, *M. musculus*, *R. norvegicus*, and *H. sapiens*) are aligned with highlighted protein domain information via collaboration with DIOPT. The key biological and genetic features are then extracted from existing model organism databases (SGD, PomBase, WormBase, FlyBase, ZFIN, MGI, and RGD).

<http://marvel.org/>

Mutations database; Coverage: Global

Matchmaker Exchange - GA4GH/IRDIRC (R)



Matchmaker Exchange allows finding matching cases with a deleterious variant in the same gene and overlapping phenotype to provide sufficient evidence to identify genetic causes for patients with rare disease. The matching of cases with similar phenotypic and genotypic profiles (matchmaking) is performed via a federated platform (Exchange) through standardized application programming interfaces (APIs) and procedural conventions.

<http://www.matchmakereexchange.org>

Data capture/analysis; Coverage: Global

Mendelian (R) **mendelian**



Search engine mapping phenotypes and genotypes, specifically designed for physicians working in Rare Disease diagnosis.

<https://www.mendelian.co>

Data capture/analysis; Coverage: Global

Metabolights (EBI) (R)



Database for Metabolomics experiments and derived information. The database is cross-species, cross-technique and covers metabolite structures and their reference spectra as well as their biological roles, locations and concentrations, and experimental data from metabolic experiments.

<https://www.ebi.ac.uk/metabolights/>

Metabolomics; Coverage: Global

Monarch Initiative (R)



Goals are to (1) Integrate, align, and re-distribute cross-species gene, genotype, variant, disease, and phenotype data; (2) Provide a portal for exploration of phenotype-based similarity; (3) Facilitate identification of animal models of human disease through phenotypic similarity; (4) Enable quantitative comparison of cross-species phenotypes; (5) Develop embeddable widgets for data exploration; (6) Influence genotype and phenotype reporting standards; (7) Improve ontologies to better curate genotype-phenotype data.

<https://monarchinitiative.org/>

Ontology; Coverage: Global

NeurOmics (Integrated European Project on Omics Research of Rare Neuromuscular and Neurodegenerative Diseases) (P)



Research consortium which brings together the leading research groups in Europe, five highly innovative SMEs as well as overseas experts in the relevant fields. Using sophisticated -omics technologies, this consortium aims to improve diagnostics and develop new treatments for ten major neurodegenerative and neuromuscular diseases affecting the cortex, basal ganglia, cerebellum, spinal cord, peripheral nerves, neuromuscular junction, and muscle.

<http://rd-neuromics.eu/>

Biomarkers; Coverage: Europe

NexProt (R) nextprot



Comprehensive human-centric discovery platform, offering its users a seamless integration of and navigation through protein-related data.

<https://www.nextprot.org>

Proteomics; Coverage: Global

NHS Clinical Research Network (National Institute for Health Research [NIHR] Clinical Research Network [CCRN]) (R)

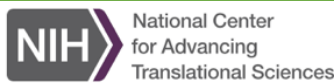


NHS CRN provides facilities and technology that enables research to thrive, funding for research studies as well as academic training, career development and research capability development.

<https://www.nihr.ac.uk>

Clinical trials; Coverage: National (UK)

NIH Rare Diseases Clinical Research Network (R)



NIH RD CRN provides support for clinical studies and facilitates collaboration, study enrolment and data sharing.

<https://ncats.nih.gov/rdcrn>

Clinical trials; Coverage: National (USA)

OMIM (Online Mendelian Inheritance in Man) (R)



Comprehensive, authoritative compendium of human genes and genetic phenotypes that is freely available and updated daily. The full-text, referenced overviews in OMIM contain information on all known mendelian disorders and over 15,000 genes.

<https://www.omim.org/>
Catalogue [diseases/genes]

ORDO (Orphanet Rare Disease Ontology) (R)



ORDO is jointly developed by Orphanet and the EBI and provides a structured vocabulary for rare diseases capturing relationships between diseases, genes and other relevant features to form a resource for the computational analysis of rare diseases.

www.orphadata.org; <https://www.ebi.ac.uk/ols/ontologies/ordo>
Ontology; Coverage: Europe

Orphamizer (R)



Orphamizer is a clinical diagnostic support tool developed using the Human Phenotype Ontology (HPO) and Orphanet data. Orphamizer helps establish a differential diagnosis by entering a combination of phenotypes using a controlled vocabulary (HPO). This beta prototype searches for semantic similarity in ontologies.

http://compbio.charite.de/phenomizer_orphanet/
Phenotype/Clinical features; Coverage: Europe

OrphanDev (R)



Public platform attached to Aix Marseille Université (AMU) and the Institut de Neurosciences de la Timone (INT). OrphanDev is specialised in supporting researchers, clinicians and the health industry actors in the development of drugs for rare diseases.

<http://orphan-dev.org/index.html>
Translational research; Coverage: National (FR)

Orphanet (R)



Resource gathering and improving knowledge on rare diseases. Orphanet maintains the Orphanet rare disease nomenclature (ORPHAnumber), essential in improving the visibility of rare diseases in health and research information systems. Comprehensive data sets related to rare diseases and orphan drugs from the Orphanet knowledge base can be accessed and downloaded in reusable formats via Orphadata.

www.orpha.net; www.orphadata.org
Catalogue [diseases/genes] ; Coverage: Europe

PatientArchive (Kinghorn Centre for Clinical Genomics (KCCG) Patient Archive platform) (R)



Secure and searchable data storage platform that allows for clinical and genomic information to be stored and tracked. It is a flexible system that allows to make file

notes after a patient enquiry or appointment, store documents and to share this information securely with other clinicians. It also allows for the storage of filtered genomic data summaries to facilitate the writing of molecular pathology reports.

<http://patientarchive.org/>

Data capture/analysis; Coverage: Global

PCBIS (Plateforme de Chimie Biologique Intégrative de Strasbourg [Platform of Integrative Chemical Biology of Strasbourg])

(R)

PCBIS gives access to High Throughput Screening (HTS), chemical libraries and target libraries to academic and private laboratories.

<http://www.pcbis.fr/fr/>

Pharmacology/Molecular screening; Coverage: National (FR)



PedCRIN (Paediatric Clinical Research Infrastructure Network)

(R)

Lunched on 1st January 2017. It brings together the European Clinical Research Infrastructure Network (ECRIN) and the founding partners of the European Paediatric Clinical Trial Research Infrastructure (EPCT-RI) to develop capacity for the management of multinational paediatric non-commercial clinical trials. PedCRIN will effectively bridge paediatricians and other partners across Europe (and internationally) to combine resources and expertise to conduct and manage robust studies, while minimising risk and protecting the child participants.

<https://www.ecrin.org/projects/pedcrin>

Clinical trials; Coverage: Europe



PhenomeCentral (R)



Repository for clinicians and scientists working in the rare disorder community. It is a centralized database where users can enter their patients' data and be connected to other patient profiles within PhenomeCentral that share similar phenotypes and genotypes.

<https://www.phenomecentral.org>

Data capture/analysis; Coverage: Global

PhenoMeNal (EBI) (R)



Comprehensive and standardised e-infrastructure that supports the data processing and analysis pipelines for molecular phenotype data generated by metabolomics applications.

<http://portal.phenomenal-h2020.eu/home>

Metabolomics; Coverage: Europe

PHENOMIN (PHENOMIN: French National Infrastructure in Biology and Health for mouse phenogenomics) (R)



PHENOMIN supports the scientific community for the best use of mouse phenogenomics in Research and is a unique resource of model animals for fundamental research, biotechnological and biopharmaceutical innovations.

<http://www.phenomin.fr/en-us/>

Models (animal, cells) ; Coverage: National (FR)

Phenomizer (R)



Software that aims to help clinicians to identify the correct differential diagnosis in the field of human genetics. The user enters the signs/symptoms of the patient encoded as terms from the Human Phenotype Ontology. The software then ranks all diseases from OMIM, Orphanet, and DECIPHER by a score that reflects how well the phenotypic profiles of the patient and the disease match to each other.

<http://compbio.charite.de/phenomizer/>

Phenotype/Clinical features; Coverage: Europe

PhenoTips (R)



Free and open-source software tool for collecting and analysing standardised phenotypic information of patients with genetic disorders. Information is collected with customised forms using vocabularies and ontologies such as the Human Phenotype Ontology (HPO), the Orphanet Rare Disease Ontology (ORDO) and Online Mendelian Inheritance in Man (OMIM).

<https://phenotips.org/>

Clinical descriptions; Coverage: Global

PPRL (IRDiRC/GA4GH) (Privacy-Preserving Record Linkage) (R)



PPRL is a joint IRDiRC-GA4GH Task Force that prepares policy and technology standards to enable highly reliable linking of coded data records associated with the same individual without disclosing the identity of that individual.

<http://www.irdirc.org/activities/task-forces/privacy-preserving-record-linkage/>

Data access/sharing; Coverage: Europe

PRIDE (PRoteomics IDentification database) (EBI) (R)



Centralized, standards compliant, public data repository for proteomics data, including protein and peptide identifications, post-translational modifications and supporting spectral evidence.

<https://www.ebi.ac.uk/pride/archive/>

Proteomics; Coverage: Global

RaDiCo (Rare Disease Cohorts) (R)



Large epidemiological instruments or « cohorts » in the field of rare diseases.

<https://radico.fr/en/accueil>

Patient cohorts; Coverage: Europe

RD-Connect (R)



Integrated platform connecting databases, registries, biobanks and clinical bioinformatics for rare disease research.

<https://rd-connect.eu>; <http://rd-connect.eu/platform/registries/>;
<http://catalogue.rd-connect.eu/>; <https://platform.rd-connect.eu>

RD-integrated resources; Coverage: Europe

RD-Connect Genome-Phenome Analysis Platform (R)



Online tool for diagnosis and gene discovery in rare disease research. The user-friendly interface lets users, even without bioinformatic training, analyse Next-Generation Sequencing data, such as genomes and exomes, linked to detailed clinical information (phenotypic data) stored in the PhenoTips database.

<https://platform.rd-connect.eu>

Data capture/analysis; Coverage: Europe

RD-Connect Registry & Biobank Finder (R)



Online directory of existing rare disease databases, registries and biobanks.

<http://catalogue.rd-connect.eu>

Catalogue [registries/biobanks]; Coverage: Europe

RD-Connect Sample Catalogue (R)



Platform where researchers can find detailed information on individual samples stored in rare disease biobanks and provides information about the available biomaterials, such as primary cells, tissue, DNA, serum, RNA, cell lines and others.

<https://samples.rd-connect.eu>

Catalogue [biobanks] ; Coverage: Europe

RD-Nexus (R)



An array of networks (e.g., ERNs) making their assets safely discoverable, without ever exposing the data content of the asset itself. Leaves all control over who can query, what is queried, and what responses they get, in the hands of the data owner.

<http://rdnexus.com>

Data access/Sharing; Coverage: Europe

Research Data Alliance (R)



Provides a neutral space where its members can come together through focused global Working and Interest Groups to develop and adopt infrastructure that promotes data-sharing and data-driven research, and accelerate the growth of a cohesive data community that integrates contributors across domain, research, national, geographical and generational boundaries.

<https://www.rd-alliance.org/about-rda/who-rda.html>

Data access/Sharing; Coverage: Global

RISCAPE (Research InfraStructure landscape) (R)



Provides a consistent report of the international research infrastructure landscape, particularly to help European Commission, other funding institutions and the ESFRI in their work supporting development of the tools for the science.

<http://www.riscape.eu/>

Translational research; Coverage: Europe

SEFALer (Servicio de Fenotipado de Animales de Laboratorio en red (Service for Network Phenotyping of Laboratory Animals)) (R)



Specialises in phenotyping animal models with specific applications for rare diseases. SEFALer is a network technology platform which research groups and services (SEFALer units) are joining to carry out work on phenotyping laboratory animals, particularly genetically modified mice.

<https://www.ciberer.es/en/platforms/sefaler>

Models (animal, cells); Coverage: National (ES)

Solve-RD (Solving the Unsolved Rare Diseases) (P)



Research project which goal is to deliver diagnostic tests for most rare diseases by 2020. Ambitions are to solve large numbers of rare disease, for which a molecular cause is not known yet by sophisticated combined omics approaches, and to

improve diagnostics of rare disease patients through contribution to, participation in and implementation of a “genetic knowledge web” which is based on shared knowledge about genes, genomic variants and phenotypes.

<http://solve-rd.eu>

Data capture/analysis; Coverage: Europe

TREAT-NMD Patient Registries (Treat NeuroMuscular Diseases) (R)



TREAT-NMD was initially established as a EU funded ‘network of excellence’ with the remit of ‘reshaping the research environment’ in the neuromuscular field. The network has developed from its European roots to become a global organization bringing together leading specialists, patient groups and industry representatives to ensure preparedness for the trials and therapies of the future while promoting best practice today.

<https://treat-nmd.org>

Catalogue [registries]; Coverage: Europe

UniProt (Universal protein knowledgebase) (R)



Provide the scientific community with a comprehensive, high-quality and freely accessible resource of protein sequence and functional information.

<http://www.uniprot.org/>

Proteomics; Coverage: Global

WGGC (West German Genome Centre) (R)



WGGC Provides the expertise to carry out research projects that are based on genomics data as well as the infrastructure necessary for sequencing, microarray analyses and other high-throughput methods in DNA and RNA analysis as well as bioinformatics.

<http://medfak.uni-koeln.de/genomik.html?&L=1>

Genomics; Coverage: National (DE)

Wormbase (R)



Centralized database for information pertaining to nematode genes and genomes.

<https://www.wormbase.org>

Genomics; Coverage: Global



ANNEX II. List of WG-identified resources contacted for the 2017 survey

NB: Resources are sorted by type. Several resources can correspond to more than one type.

Type	Name	URL
Biobanks	ADOPT-BBMRI	http://www.bbmri-eric.eu/scientific-collaboration/adopt-bbmri-eric/
Biobanks	BBMRI	www.bbmri.eu
Biobanks	EuroBioBank	http://www.eurobiobank.org
Catalogues	ORPHANET	http://www.orpha.net/ ; www.orphadata.org
Catalogues	RD-CONNECT	http://rd-connect.eu/platform/registries/ ; http://catalogue.rd-connect.eu/
Catalogues	Treat-NMD	http://www.treat-nmd.eu/
Clinical research	ECRIN	http://www.ecriin.org
Clinical research	EORTC	www.eortc.org
Clinical research	GRIP (Global research in Paediatrics)	http://www.grip-network.org/
Clinical research	IDEAL	http://www.ideal.rwth-aachen.de/
Clinical research	INSPIRE	https://www2.warwick.ac.uk/fac/med/research/hscience/stats/currentprojects/inspire/
Clinical research	PED-CRIN	http://www.ecriin.org/news/ecriin-receives-h2020-funding-pedcrin-paediatric-network-0
Data sharing/ analysis	ADA-M Matrix	https://genomicsandhealth.org/work-products-demonstration-projects/automatable-discovery-and-access-matrix

Type	Name	URL
Data sharing/ analysis	Beacon	https://genomicsandhealth.org/work-products-demonstration-projects/beacon-project-0
Data sharing/ analysis	Elixir	https://www.elixir-europe.org/
Data sharing/ analysis	EOSC	https://eoscpilot.eu/
Data sharing/ analysis	EUPID	https://eupid.eu
Data sharing/ analysis	PPRL	http://www.irdirc.org/activities/current-activities/privacy-preserving-record-linkage/
Data sharing/ analysis	RD-CONNECT	http://rd-connect.eu/platform/registries/ ; http://catalogue.rd-connect.eu/
Data sharing/ analysis	RDA-Bio-sharing	https://fairsharing.org/
Diagnostic methods	Orphanizer	http://compbio.charite.de/pheno_mizer_orphanet
Drug discovery	ChemBioNet	www.chembionet.info
Drug discovery	CIDSTEM	http://www.cidstem.unimore.it/chiamo.html
Drug discovery	EU-OPENSREEN	http://www.eu-openscreen.eu/
Drug discovery	I-STEM	http://www.eurostemcell.org/
Genotype/ Phenotype	DECIPHER	https://decipher.sanger.ac.uk/
Genotype/ Phenotype	EGA	https://www.ebi.ac.uk/ega/home
Genotype/ Phenotype	LOVD	http://www.lovd.nl
Imaging	EuroBioImaging	http://www.eurobioimaging.eu/

Type	Name	URL
Infrastructures	CORBEL	http://www.corbel-project.eu/home.html
Models	CIDSTEM	http://www.cidstem.unimore.it/chi_siamo.html
Models	I-STEM	http://www.eurostemcell.org/
Models	INFRAFRONTIER	https://www.infrafrontier.eu
Models	Wormbase	http://www.wormbase.org/#012-34-5
Omics	COSMOS	http://www.cosmos-fp7.eu/
Omics	EuRenOmics	https://www.eurenomics.eu/
Omics	NeurOmics	http://rd-neuromics.eu/
Omics	PhenoMeNal	http://portal.phenomenal-h2020.eu/home
Omics	UniProt	http://www.uniprot.org/
Ontologies/ Standards	HGNC	http://www.genenames.org/
Ontologies/ Standards	HPO	http://human-phenotype-ontology.github.io/about.html
Ontologies/ Standards	ORPHANET	http://www.orpha.net/; www.orphadata.org
Registries/ cohorts	JRC	https://ec.europa.eu/jrc/en
Registries/ cohorts	RaDiCo	www.radico.fr

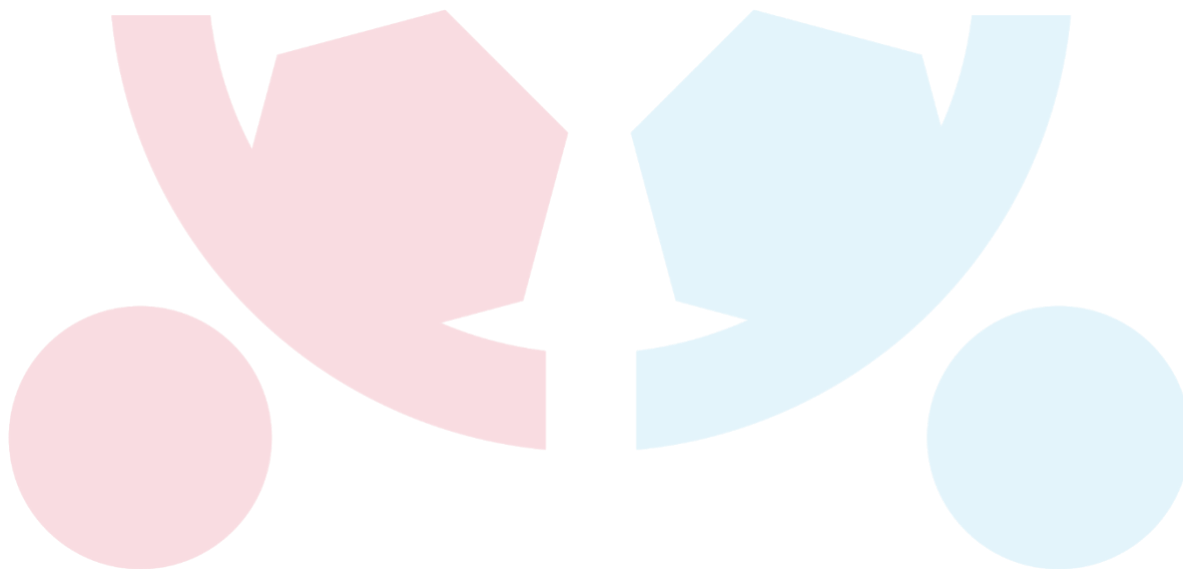
ANNEX III. List of questions asked in the 2017 survey

Please describe the main goals of your project/resource
Please indicate the starting and the end date of your project
<p>Compared to your initial goals, would you say:</p> <p>Option 1: We have achieved 100% of our goals, Option 2: We have delivered tools/results that are fully operational</p> <p>Option 2: We have delivered tools/results that are fully operational, Option 3: We have delivered tools/results that need more development/investment to be fully operational</p> <p>Option 3: We have delivered tools/results that need more development/investment to be fully operational</p> <p>Option 4: We are still in the development phase</p>
<p>Are you conducting activities specifically related to the RD field?</p> <p>Option 1: Yes, all our activities are intended to RD</p> <p>Option 2: Yes, some of our activities are related to RD</p> <p>Option 3: No, but our generic tools/results could be adapted/useful for RD</p>
Are developments (tools, methods...) specifically intended for RD planned in your future strategy?
Is your resource IRDiRC-recognized or does it fulfil the criteria fully or partially (http://www.irdirc.org/activities/irdirc-recognized-resources/)?
Are you involved in the development of new therapies for RD?
<p>Are you involved/developing tools for:</p> <ul style="list-style-type: none"> - Drug development - Drug discovery - Interpretation of variants and sharing patient data - Variant mapping - Linking data records of individuals in a Privacy Preserving setting - Providing background genomic data useful for gene-specific therapy development - Signposting treatable causative variants and potential therapies (future development) - The design/conduct of RD clinical trials - The identification of Biomarkers for RDs - The evaluation of the safety/efficacy and SAEs of therapies for RDs - The validation of e-health tools and connected devices
Has your project/resource a direct impact on RD research and RD patients (diagnosis or treatment)?

Please provide PubMed references that show the impact of your resource/infrastructure on RDs
<p>Are your tools/results interoperable?</p> <ul style="list-style-type: none"> - Option 1: Yes, SEMANTIC interoperability has been foreseen in the project. - Option 2: Yes, TECHNICAL interoperability has been foreseen in the project.
<p>Are the licensing issues/ terms and conditions for re-use of your tools/results clearly established?</p>
<p>Who holds the Intellectual Property Rights of your results/tools? (Who is the owner)</p>
<p>Is your project sustainable?</p> <ul style="list-style-type: none"> - Option 1: Yes, fully sustainable - Option 2: Yes, partially sustainable - Option 3: No, no further funding is envisioned after the end of the project
<p>Is your resource scalable?</p>
<p>Are patients/patient organisations involved in your project?</p>
<p>Who are your users?</p>
<p>Who are your partners?</p>
<p>Has your resource/infrastructure already collaborated with or made concrete plans for collaboration with European Reference Networks (ERNs)?</p>
<p>If you were an EJP-RD partner, what would be your potential contribution?</p>
<p>What would you expect from participating to EJP-RD?</p>
<p>Would you accept to be contacted for further discussions?</p>

ANNEX IV. List of questions asked in the 2019 survey

What is the name of you resource?
What is the mission of your resource?
Please give a brief summary of your resource?
Why is your resource relevant to rare diseases?
Is your resource RD-specific?
What kind of data does the resource deal with?
Which standards, ontologies, vocabularies are already used by your resource?
What is the access model?
To what extent is the resource FAIR? Does it have accessible APIs?
Is the resource IRDiRC-recognized or does it fulfil the criteria fully or partially (http://www.irdirc.org/activities/irdirc-recognized-resources/)?
Indicate in which projects/initiatives your resource is involved in and their main aim (1 line for each).
What are the planned contributions from the EJP to/for the resource, especially in years 1 and 2?
What are the planned contributions from the resource to the EJP, especially in years 1 and 2? (adaptations, development needs...)
Any other comments?



ANNEX V. Results of the 2017 survey on RD resources

Resources	Maturity (achievements vs goals)	Involvement on RD	Plans for further/increased involvement on RD	Interoperability	Transparency of terms of use/IP issues	Sustainability
ChemBioNet	Yes	No	Possible	Technical	Yes	Partially
DECIPHER	Yes	Yes	Yes	Technical & Semantic	Yes	Fully
EATRIS-ERIC	partially	Yes	Yes	No	Yes	Fully
EORTC-SPECTA	Yes	Yes	Yes	Technical & Semantic	Yes	Fully
EU-OPENSREEN	No	No	Possible	No	Yes	Fully
Euro-Biolmaging	No	No	Possible	Technical	No	Partially
EuroBioBank	Partially	Yes	Yes	Semantic	Yes	Partially
ECRIN	Partially	Yes	Yes	Technical & Semantic	Yes	Fully
EU RD Platform	Partially	Yes	yes	Technical & Semantic	Yes	Fully
GO-FAIR	Partially	Yes	Yes	Technical & Semantic	Yes	Partially
FAIRsharing	Partially	No	Possible	No	Yes	Partially
I-STEM	Yes	Yes	Yes	No	Yes	Fully
IDEAL	Yes	Yes	Yes	No	Yes	No
Orphanet and derived	Yes	Yes	Yes	Technical & Semantic	Yes	Partially
PEDCRIN	No	No	Yes	Technical	No	Partially
RaDiCo	Yes	Yes	yes	Technical & Semantic	Yes	Partially
RD-Connect	Partially	Yes	Yes	Technical & Semantic	No	Partially
SOLVE-RD	No	Yes	Yes	Technical & Semantic	No	No
UniProtKB	Yes	Yes	Yes	Technical & Semantic	Yes	Fully

Resources	Users, etc	IRDiRC recognized	Involvement of PO	Collaboration / plans with ERNs	Input to EJP	Output from EJP
ChemBioNet	Needs further information	No	No	No	Tools/Platform; Personnel	Networking, visibility, funding
DECIPHER	Sufficient information	Yes	No	No	Tools/Platform	Networking, visibility
EATRIS-ERIC	Needs further information	No	Yes	Yes	Strategic contribution	Networking, visibility, funding
EORTC-SPECTA	Sufficient information	No	Yes	Yes	Tools/Platform; Personnel	Networking, visibility, funding
EU-OPENSREEN	Needs further information	No	No	No	Tools/Platform	Networking, visibility
Euro-BioImaging	Needs further information	No	No	No	Tools/Platform	Networking, visibility
EuroBioBank	Sufficient information	No	Yes	Yes	Tools/Platform; Personnel	Networking, visibility, funding
ECRIN	Needs further information	No	Yes	No	Tools/Platform	Networking
EU RD Platform	Needs further information	No	No	Yes	Tools/Platform	Networking, visibility
GO-FAIR	Needs further information	No	Yes	Yes	Tools/Platform; Personnel	Networking, visibility, funding
FAIRsharing	Sufficient information	No	No	No	Tools/Platform	Networking, visibility, funding
I-STEM	Needs further information	No	Yes	No	Tools/Platform; Personnel	Networking, visibility, funding
IDEAL	Needs further information	No	Yes	No	N.A.	N.A.
Orphanet and derived	Sufficient information	Yes	Yes	Yes	Tools/Platform; Personnel	Networking, visibility, funding
PEDCRIN	Needs further information	No	Yes	No	Tools/Platform; Personnel	Networking, visibility, funding
RaDiCo	Needs further information	No	Yes	Yes	Tools/Platform; Personnel	Networking, visibility, funding
RD-Connect	Needs further information	No	Yes	Yes	Tools/Platform; Personnel	Networking, visibility, funding
SOLVE-RD	Needs further information	No	Yes	Yes	Tools/Platforms	Networking, visibility, funding
UniProtKB	Sufficient information	No	No	No	Tools/Platform; Personnel	Networking, visibility, funding

ANNEX VI. Results of the 2019 survey on RD resources

<p style="text-align: center;">BBMRI-ERIC</p> <p>RD-specific: No Type of data: Catalogue [biobanks] Standards, ontologies, vocabularies used in resource: MIABIS, OBIB, ICD-10, ORDO, SNOMED CT (and many others). Access model: Variable Resource is FAIR: Yes IRDiRC-recognized: No Involvement in other projects/initiatives: • EOSC-Life - EOSC for Life Sciences RIs • EJP RD • ISO - standardization of quality • BBMRI-ERIC is heavily involved in many projects - see http://www.bbmri-eric.eu/eu-grants/. Output from EJP: Networking Input to EJP: Tools/Platform</p>
<p style="text-align: center;">Cellosaurus</p> <p>RD-specific: No Type of data: Models [animals] Standards, ontologies, vocabularies used in resource: • For disease terms: NCI thesaurus • For organisms: NCBI taxonomy • For chemicals: EBI ChEBI; DrugBank • For genes: Human: HGNC; Mouse: MGI; Rat: RGD; Vertebrates: VGNC; Drosophila: FlyBase • For proteins: UniProtKB • Sequence variations: HGVS nomenclature • For STR markers: ANSI/TCC ASN-0002-2011 + additional markers • Other in house small "vocabularies": cell line categories, MHC genes, Ig isotypes, genders. Access model: Open access Resource is FAIR: No IRDiRC-recognized: No (rejected) Involvement in other projects/initiatives: • RRI: The Resource Identification • ICLAC: International Cell Line Authentication Committee • hPSCreg. Output from EJP: Networking; Visibility Input to EJP: Tools/Platform</p>
<p style="text-align: center;">DECIPHER</p> <p>RD-specific: Yes Type of data: Phenotype/Genomics Standards, ontologies, vocabularies used in resource: HPO, OMIM disease terms, HGVS, ACMG recommended vocabulary. Access model: Open access Resource is FAIR: Yes IRDiRC-recognized: Yes Involvement in other projects/initiatives: • GA4GH, includes MME interaction and involvement with GA4GH Case Discovery API • The Transforming Genetic Medicine Initiative (TGMI) • ESHG/Eurogentest two-dimensional variant classification system. Output from EJP: Networking; Guidance Input to EJP: Tools/Platform</p>

Ensembl Variant Effect Predictor (VEP)

RD-specific: No

Type of data: Genomic

Standards, ontologies, vocabularies used in resource: SO, HPO, EFO, ORDO.

Access model: Open access

Resource is FAIR: Yes

IRDiRC-recognized: No

Involvement in other projects/initiatives: • WP11 Collaborative integrated analysis of multi-omics data from selected use cases & User-driven implementation of new analysis functionalities and tools • WP13 Mapping variants to function & Mapping to genes.

Output from EJP: None

Input to EJP: Tools/Platform

EuroBioBank

RD-specific: Yes

Type of data: Catalogue [biobanks/registries]

Standards, ontologies, vocabularies used in resource: Ontologies

Access model: Open access

Resource is FAIR: Yes

IRDiRC-recognized: No

Involvement in other projects/initiatives: • TREAT-NMD • RD-CONNECT • BBMRI • Several individual projects where the samples are used.

Output from EJP: Networking

Input to EJP: Tools/Platform

HPO

RD-specific: No

Type of data: Phenotypes/Clinical features

Standards, ontologies, vocabularies used in resource: MONDO, GO, ChEBI, Uberon, others.

Access model: Open access

Resource is FAIR: Yes

IRDiRC-recognized: Yes

Involvement in other projects/initiatives: • Solve-RD • Monarch initiative • GA4GH • Orphanet.

Output from EJP: Networking

Input to EJP: Tools/Platform

hPSCreg

RD-specific: No

Type of data: Models [cells]

Standards, ontologies, vocabularies used in resource: hPSCreg is using EBI Biosamples IDs for samples and donors, EBI ZOOMA for mapping free text to ontology terms and EBI Ontology Lookup Service.

Access model: Open access

Resource is FAIR: Yes

IRDiRC-recognized: No

Involvement in other projects/initiatives: • EBISC (<https://ebisc.org/>) • International Stem Cell Banking Initiative (ISCBI) • Global Alliance of iPSC Therapy (GAiT).

Output from EJP: Networking; Visibility

Input to EJP: Tools/Platform

INFRAFRONTIER

RD-specific: No

Type of data: Models [animals]

Standards, ontologies, vocabularies used in resource: We follow the rules and guidelines established by the International Committee on Standardized Genetic Nomenclature for Mice, use MGI (Mouse Genome Informatics) gene and allele IDs, Mouse Phenotype Ontology (MP), Diseases Ontology (DO) and Orphanet/ORDO.

Access model: Service fee

Resource is FAIR: No

IRDiRC-recognized: No

Involvement in other projects/initiatives: • International Mouse Phenotyping Consortium (IMPC) • CORBEL • EOSC-Life • EJP-RD • Pathbio ERASMUS + Knowledge Alliance consortium.

Output from EJP: Networking; Guidance

Input to EJP: Tools/Platform

MetaboLights

RD-specific: No

Type of data: Metabolomics

Standards, ontologies, vocabularies used in resource: MetaboLights employs the ZOOMA ontology annotation tool to submit and map terms to the most relevant available ontologies. Examples include (but are not limited to); Coition, Experimental factor ontology (EFO), Orphanet, Chemical Entities of Biological Interest (ChEBI).

Metabolomics community standards (MSI) (these have overlaps with the GSC standards and we are actively discussing collaboration between them to ensure interoperability).

Access model: Open access

Resource is FAIR: Yes

IRDiRC-recognized: No

Involvement in other projects/initiatives: • Metabolomics Quality Assurance and quality Control Consortium (mQACC) • The OECD Metabolomics Reporting Framework (MRF) • PhenoMeNal • The METASPACE platform.

Output from EJP: Networking; Guidance

Input to EJP: Tools/Platform

neXtProt

RD-specific: No

Type of data: Proteins

Standards, ontologies, vocabularies used in resource: NCI thesaurus, UniProtKB, EBI ChEBI, DrugBank, HGNC, HGVS, CALOHA (in house developed ontology with links to BRENDA, Uberon), Cellosaurus, GO annotations, Mammalian phenotype ontology, ICEPO (in house developed ontology), Protein Property (in house developed ontology).

Access model: Open access

Resource is FAIR: Yes

IRDiRC-recognized: No

Involvement in other projects/initiatives: • HUPO HPP

Output from EJP: None

Input to EJP: Tools/Platform

Orphanet, Orphadata, ORDO, HOOM

RD-specific: Yes

Type of data: Catalogue

Standards, ontologies, vocabularies used in resource: Orphanet produces the standard nomenclature and ontology for rare diseases; it produces mappings to other standards, terminologies and resources: HPO; ICD10; ICD11; SNOMED CT; OMIM; GARD; HGNC; MedDRA; UMLS; ICF-YA; the list being evolutive.

Access model: Open access

Resource is FAIR: Yes

IRDiRC-recognized: Yes

Involvement in other projects/initiatives: • RD-CODE • SOLVE-RD • Rare2030 • EJP-RD • HIPBI-RD • OrphaNetWork • JARC • RD-ACTION.

Output from EJP: Networking; Guidance

Input to EJP: Tools/Platform

PRIDE

RD-specific: No

Type of data: Proteomics

Standards, ontologies, vocabularies used in resource: We support the relevant mass spectrometry proteomics data standards produced by the Proteomics Standards Initiative (<http://www.psiview.info/>), including mzML, mzIdentML, mzTab, among others. In terms of controlled vocabularies, we use specifically PSI-MS (for mass spectrometry related information), PSI-MOD and UniMod (for post-translational modifications) and others such as NCBI Taxonomy, EFO, BRENDA, and the Cell ontology.

Access model: Open access

Resource is FAIR: Yes

IRDiRC-recognized: No (but would fulfil criteria)

Involvement in other projects/initiatives: • EGA and MetaboLights • Many more...

Output from EJP: _

Input to EJP: _

RaDiCo – Rare Disease Cohorts

RD-specific: Yes

Type of data: Catalogue

Standards, ontologies, vocabularies used in resource: • Human Phenotype Ontology • MedDRA • Thériaque. RaDiCo platform is also implementing WHO Drug Global, the international reference for medicinal product information.

Access model: Double authentication process.

Resource is FAIR: No

IRDiRC-recognized: No

Involvement in other projects/initiatives: • Several Public Private Partnerships • National RD Program (PNMR3) • EJP RD.

Output from EJP: Networking

Input to EJP: Tools/Platform

RD-Connect GPAP

RD-specific: Yes

Type of data: Phenotype/Genomics/Transcriptomics

Standards, ontologies, vocabularies used in resource: HPO, OMIM, ORDO, HGNC.

Access model: Data Access Committee

Resource is FAIR: Yes

IRDiRC-recognized: Yes

Involvement in other projects/initiatives: • Solve-RD

Output from EJP: Networking

Input to EJP: Tools/Platform

RD-Connect Sample Catalogue

RD-specific: Yes

Type of data: Catalogue [biosamples]

Standards, ontologies, vocabularies used in resource: ORDO, FMA, HPO, UBERON, OBIB (MIABIS).

Access model: Open access

Resource is FAIR: Yes

IRDiRC-recognized: No

Involvement in other projects/initiatives: • RD-Connect • BBMRI-ERIC • EJP-RD.

Output from EJP: Networking

Input to EJP: Tools/Platform

RD-NEXUS (based on Cafe Variome)

RD-specific: Yes

Type of data: Catalogue

Standards, ontologies, vocabularies used in resource: Any and all recommended by GA4GH and Orphanet

Access model: Variable

Resource is FAIR: Findable

IRDiRC-recognized: No (but would fulfil criteria)

Involvement in other projects/initiatives: • Current EU/IMI projects EPAD • Solve-RD • EJP-RD.

Output from EJP: Networking

Input to EJP: Tools/Platform

The European Platform on Rare Disease Registration (EU RD Platform)

RD-specific: Yes

Type of data: Catalogue

Standards, ontologies, vocabularies used in resource: Orphanet classification of diseases, ICD-10, ICD9 Orpha code, Alpha code, International classification of mutations (HGVS), HGNC, OMIM code, international Classification of Functioning and Disability (ICF).

Access model: Open access

Resource is FAIR: Yes

IRDiRC-recognized: No

Involvement in other projects/initiatives: • European Commission's initiatives related to rare diseases

Output from EJP: _

Input to EJP: Tools/Platform

UniProt

RD-specific: No

Type of data: Proteins

Standards, ontologies, vocabularies used in resource: International protein/gene nomenclature guidelines, sequence download formats (Fasta, PEFF, XML, RDF etc), ChEBI/RHEA/EC, imports and displays GO annotations, internal published CVs.

Access model: Open access

Resource is FAIR: Yes

IRDiRC-recognized: No

Involvement in other projects/initiatives: _

Output from EJP: Networking

Input to EJP: Tools/Platform

ANNEX VII. List of members of the 2017 working group

Name	Institute	Country	Platform/Network	Main field of expertise
Peter Antal	University of Technology	Hungary		RD expert
Kaan Boztug	Ludwig Boltzmann Institute of Rare and Undiagnosed Diseases, Vienna	Austria		RD expert
Johannes Bauer	University Clinic for Dermatology of the SALK/Paracelsus Medical University	Austria	ERN Skin	ERN- Rare skin diseases
Stephane Lejeune	EORTC	Belgium	EORTC	Rare cancers, Clinical research
Denis Lacombe	EORTC	Belgium	EORTC	Rare cancers, Clinical research
Barisic Ingeborg	Children's university hospital of Zagreb	Croatia	Orphanet Croatia, EUROCAT	registries, database
Anna Skakkebaek	Aarhus University Hospital	Danemark		RD expert
Mette Moller Handrup	Aarhus University Hospital	Danemark		RD expert
Anna-Elina Lehesjoki	Neuroscience Center and Folkhälsan Institute of Genetics Neuroscience	Finland		RD expert
Estelle Gandjbakhch	APHP	France	ERN GUARD-HEART	ERN- Rare cardiac diseases
Jean Donnadieu	APHP	France	ERN EuroBlooNet	ERN- Rare hematological diseases
Xavier Jeunemaitre	APHP/HEGP	France	VASCERN	ERN- Rare vascular diseases
Marine Hurard	APHP/HEGP	France	VASCERN	ERN- Rare vascular diseases

Name	Institute	Country	Platform/Network	Main field of expertise
Jean-Yves Blay	CHU de Lyon	France	EURACAN	ERN- Rare adult cancer
Jacques Demotes	ECRIN	France	ECRIN	Clinical trials
Christine Kubiak	ECRIN	France	ECRIN	Clinical trials
Virginie Bros-Facer	EURORDIS	France	EURORDIS	Patient organizations
Ana Rath	Inserm	France	Orphanet, Solve-RD	Database, ontology
Nicolas Garcelon	Institut Imagine	France	Institut Imagine	Bioinformatics
Vincent Benoît	Institut Imagine	France	Institut Imagine	Bioinformatics
Rima Nabbout	Necker Hospital, Imagine institute	France	EPICARE (ERN)	ERN- Rare epilepsies, clinical research
Catherine Champseix	Necker Hospital	France	ERN Skin	ERN- Rare skin diseases
Christine Bodemer	Necker Hospital	France	ERN Skin	ERN- Rare skin diseases
Sophie Bernichtein	UPMC	France	ERN-RND	ERN- Rare neurological diseases
Olaf Riess	University of Tubingen	Germany	NeurOmics, Solve-RD	Omics, Solve-RD
Holm Graessner	University of Tübingen	Germany	ERN-RND, Solve-RD	ERN- Rare neurological diseases
Eileen Tracy	Mater Misericordiae University Hospital	Ireland	Orphanet Ireland	registries, database
Lucia Monaco	Fondazione Telethon	Italy	RD-Connect; TREAT-NMD	Biobanks, databases, clinical trials, innovative therapies, public-

Name	Institute	Country	Platform/Network	Main field of expertise
				private partnerships
Maurizio Scarpa	Helios Klinik Wiesbaden, University of Padova	Italy	MetabERN (ERN)	ERN- Rare metabolic diseases, clinical research
Luca Sangiorgi	Instituto Ortopedico Rizzoli, Bologna	Italy	ERN BOND	ERN- Rare bone diseases
Bruno Dallapiccola	Paediatric hospital "Bambino Gesù"	Italy	Orphanet Italy	RD expert
Anna Villa		Italy		RD expert
Sandra Katarina Alves	Portuguese Ministry of Health (National Health Institute Ricardo Jorge)	Portugal		RD expert
Pablo Daniel Lapunzina Badia	CIBERER	Spain	Rare Diseases Networking Biomedical Research Centre	RD research
Ivo Gut	CNAG, Spain	Spain	RD-Connect, Elixir	Platforms, Next-generation sequencing, bioinformatics
Manuel Posada	ISCIII/Institute of Rare Disease Research	Spain	RD Connect	Registries
Sergi Beltran	National Genomics Centre CNAG, Barcelona	Spain	RD-Connect, Solve-RD	Next-generation sequencing, bioinformatics, MatchMaker
Anna Lindstrand	Karolinska Institute	Sweden		RD expert
Saskia Karg	University of Zurich	Switzerland	Rare Diseases Initiative Radiz, CH	RD expert
Kristina Hettne	LUMC	The Netherlands	RD-Connect	Data sharing

Name	Institute	Country	Platform/Network	Main field of expertise
Marco Roos	LUMC	The Netherlands	RD-Connect, Elixir	Data sharing
Marjon Pasmooij	University Medical Center of Groningen	The Netherlands		RD expert
Helen Cross	UCL Institute of Child Health	United Kingdom	EPICARE	ERN- Rare epilepsies
Mark Turner	University of Liverpool	United Kingdom	PedCRIN	Pediatric clinical trials
Rachel Thompson	University of Newcastle	United Kingdom	RD-Connect	Registries, Biobanks, platforms
Hanns Lochmuller	University of Newcastle	United Kingdom	RD-Connect, ELIXIR	Registries, Biobanks, platforms