RESEARCH

Open Access

The RaDiCo information system for rare disease cohorts



Paul Landais^{1*}, Sonia Gueguen², Annick Clement³, Serge Amselem⁴ and for the RaDiCo Program

Abstract

Background Rare diseases (RDs) clinical care and research face several challenges. Patients are dispersed over large geographic areas, their number per disease is limited, just like the number of researchers involved. Current databases as well as biological collections, when existing, are generally local, of modest size, incomplete, of uneven quality, heterogeneous in format and content, and rarely accessible or standardised to support interoperability. Most disease phenotypes are complex corresponding to multi-systemic conditions, with insufficient interdisciplinary cooperation. Thus emerged the need to generate, within a coordinated, mutualised, secure and interoperable framework, high-quality data from national or international RD cohorts, based on deep phenotyping, including molecular analysis data, notably genotypic. The RaDiCo program objective was to create, under the umbrella of Inserm, a national operational platform dedicated to the development of RD e-cohorts. Its Information System (IS) is presented here.

Material and methods Constructed on the cloud computing principle, the RaDiCo platform was designed to promote mutualization and factorization of processes and services, for both clinical epidemiology support and IS. RaDiCo IS is based on an interoperability framework combining a unique RD identifier, data standardisation, FAIR principles, data exchange flows/processes and data security principles compliant with the European GDPR.

Results RaDiCo IS favours a secure, open-source web application in order to implement and manage online databases and give patients themselves the opportunity to collect their data. It ensures a continuous monitoring of data quality and consistency over time. RaDiCo IS proved to be efficient, currently hosting 13 e-cohorts, covering 67 distinct RDs. As of April 2024, 8063 patients were recruited from 180 specialised RD sites spread across the national territory.

Discussion The RaDiCo operational platform is equivalent to a national infrastructure. Its IS enables RD e-cohorts to be developed on a shared platform with no limit on size or number. Compliant with the GDPR, it is compatible with the French National Health Data Hub and can be extended to the RDs European Reference Networks (ERNs).

Conclusion RaDiCo provides a robust IS, compatible with the French Data Hub and RDs ERNs, integrated on a RD platform that enables e-cohorts creation, monitoring and analysis.

Keywords Information system, Cloud computing, Infrastructure as a Service, Interoperability, GDPR, E-cohorts, Rare diseases, French Data Hub, European Research Network

*Correspondence: Paul Landais paul.landais@umontpellier.fr Full list of author information is available at the end of the article



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Background and significance

A disease is said to be rare when it affects fewer than 1 in 2000 individuals. Rare diseases (RDs) affect 26–30 million people in Europe and at least 3 million people in France. Around 70% of RDs begin in childhood [1]. RDs are often progressive, degenerative, and disabling, accounting for over a third of all deaths occurring in childhood. There is no treatment for most of them. They have a significant impact on patients' quality of life, as well as on healthcare systems. RDs are mainly of genetic origin. Currently, 4913 genes have been implicated in 7532 different disease conditions [2].

In France, three consecutive national RD plans (i.e., "PNMR" 1, 2 and 3, standing for Plan National Maladies Rares 1, 2 and 3) have structured the healthcare offering [3]. They rely on 23 RD Healthcare Networks (RDHNs), whose topics are like those covered by the 23 ERNs, and represent 2241 accredited sites throughout the country. A National Rare Disease Data Bank [4], the BNDMR [5], registers RD patients based on a unique RD identifier [6] and a minimum data set [7]. Thus, nearly 1.5 million patients and 5700 distinct diagnoses have already been identified.

RDs clinical care and research face several challenges. The number of patients per disease is limited. Patients are dispersed over large geographic areas, making data collection difficult. For a given RD, the number of researchers involved is very limited. Current databases as well as biological collections, when existing, are generally local, of modest size, incomplete, of uneven quality, heterogeneous in format and content, and rarely accessible or standardised to support interoperability. For many RDs, the disease phenotypes are complex and correspond to multi-systemic conditions; they are partially described over time, with insufficient interdisciplinary cooperation.

Thus emerged the need to generate, within a coordinated, mutualised, secure and interoperable framework, high-quality data from national or international RD cohorts, based on deep phenotyping and including molecular analysis data, notably genotypic. The RaDiCo program, coordinated by Inserm, has been funded by the National Research Agency for the 2011–2024 period in the framework of the "Investments for the Future" program. It was dedicated to creating of a national operational platform to support RD e-cohorts selected from a national call.

To support the RD cohorts electronically, the RaDiCo platform was built with the complementary expertise of two teams: Clinical Epidemiology Research, and e-Health & Information Technology, respectively. In a previous article, we described the data related to the secondary objectives of the cohorts [8]. Here, we describe the RaDiCo IS and its specifications to support epidemiological research on RD e-cohorts.

Materials and methods

To develop the RaDiCo IS, we first defined several principles and steps (Table 1).

To achieve semantically interoperable data exchange, we used controlled terminology standards, the Orphanet terminology for RDs [9] and the Medical Dictionary for Regulatory Activities [10]. To support the interoperability, we standardised the data and used ontologies, Human Phenotype ontology [11] and Orphanet Rare Disease Ontology [12]. Exchange format standards specify the way information is encoded. We used SAS statistical software, R formats and Extensible Markup Language (xml). The standards of the Clinical Data Interchange Standards Consortium (CDISC) [13] were proposed by REDCap [14]. We also complied to the FAIR principles [15]. The cohorts' implementation followed the STROBE recommendations [16].

The RD ecosystem

RaDiCo operates in a complex ecosystem with a large number of actors. For its 13 cohorts, RaDiCo is linked to 180 RDs centers, due to the low number of patients per RD and the geographic dispersion of patients throughout France. Figure 1 describes the actors involved in the program. Identifying the plurality of the protagonists as well

Table 1 RaDiCo IS implementation steps

- 1) Identifying the RD ecosystem and related actors
- 2) Designing the functional architecture allowing access to the platform
- 3) Capturing medical data through an electronic Clinical Report Form (e-CRF)
- 4) Developing an Infrastructure as a Service (laaS) in the cloud
- 5) Conceiving a « secure by design» infrastructure
- 6) Developing applications that will allow data exchange and security
- 7) Developing security procedures compliant with the European General Data Protection Regulation (GDPR)
- 8) Conceiving the global data architecture, which articulates its different components



Fig. 1 RaDiCo and its Rare Disease ecosystem. (LEEM: The drug companies; ARIIS: Health Industry Alliance for Research and Innovation)

as their job specificities enables defining the structure of the information demand, as well as the IS needs and the interconnections to develop.

Management of the RaDiCo program

The overall management of the RaDiCo program is provided by an executive committee, a scientific committee, and an institutional committee. In addition, each cohort has its own governance structure, with its steering and scientific committees.

Before their implementation, cohort projects were all submitted to ethical and regulatory authorities and to the French Data Protection Agency. The RaDiCo executive committee that organises a mutualised administrative monitoring, ensures contacts with the National Research Agency, Inserm, ministerial authorities, universities, and industrial partners. Cohort implementation and development are planned on a cohort-by-cohort basis; this includes cohort design, identification of key milestones, implementation, and monitoring meetings. Over 50% of cohort funding is provided by industrial and academic contracts. Human resources management is carried out centrally.

Results

Drafting the e-cohorts Case Report Forms (CRFs)

The research questions for a given cohort are relayed by the eCRF produced using the REDCap software [17]. A Data Monitoring Plan integrates the monitoring checklist and defines the modalities and frequency of visits, the nature of the data and the percentage of files to be monitored. Thereafter, a Data Validation Plan is elaborated to organise and prepare the elements required for data management. The Data Management Plan is then designed to check whether the cohort is adequately meeting its objectives: compliance with the monitoring schedule, number of patients recruited, data completion rate, number of data verified, compliance with defined Key Performance Indicators (KPIs). A RACI (Responsible, Accountable, Consulted, Informed) matrix of the roles and tasks of each RaDiCo team member is summarised in Table 2. A quality committee meets every two weeks to monitor compliance with the platform's standardized operating procedures (SOPs).

Translating the clinical research needs in a capability map.

Applications designed to meet the needs of clinical research are summarised in the RaDiCo capability map (Fig. 2). This map characterises the progress of the clinical study course, the monitoring and reporting phase, the statistical analyses and reporting of results, the management of patients and professionals' data, of data repositories, access rights and security.

Planning the design of the information system (IS)

According to this map, the stages in building the IS were thus designed as follows:

- Definition of the IS needs.
- Drafting of IS specifications.
- IS development including the integration of the eCRF into the IS.

Task roles	CRF	Annotated CRF	e-CRF conception	Entry guide	e-CRF testing	Monitoring plan	Validation plan	Data- management plan	Quality steering committee
PI	С	1	I	С	С	С	С		
O&S Director	А	А	А	I	1	А	А	А	А
Clin Res PM	R	С	С	R	R	1	1	R	R
E-Health PM	С	С	С	С	С	С	С	R	R
IS Manager	1	1	1	I	1	С	С	L	I
CRTs	С	1	1	С	С		С		R
CRAs	С	1	1	С	С	С			R
Data-manager	С	R	R	А	А	R	R	R	R
Biostatistician	С	С	С	1	С	1	1	С	

Table 2 RACI matrix of the different steps of e-CRF implementation

R: Responsible; A: Accountable; C: Consulted; I: Informed, PI: Principal Investigator; O&S Director: Operational and Scientific Director; Clin Res PM: Clinical Research Project Manager; E-Health PM: Electronic-health Project Manager; IS Manager: Information System Manager; CRTs: Clinical Research Technicians; CRAs: Clinical Research Assistants

- Preparation of methods and tools.
- IS validation.
- Tests and recipes.
- Production of the first version of the IS, and thereafter of its versioning, stored in Github.

Software architecture and applications *Functionalities and users*

The RaDiCo IS allows the management of data repositories, the collection of patient data (personal and medical) and the performance of studies linked to the data collected. It hosts data linked to two types of profiles, i.e., users and patients. On the one hand, IS users are either healthcare personnel, research professionals or computer scientists, involved in data collection, formatting and processing, as well as in IS maintenance. On the other hand, the patients represent the source of data collection and may also enter themselves questionnaires such as those on quality of life or data through medical devices, if any.

The RaDiCo IS has been conceived on the principle of an Infrastructure as a Service (IaaS), a cloud computing model. The platform manages servers' middleware, application software (executables, settings), integration of a service-oriented architecture (SOA), and databases. The cloud provider, either Microsoft Azure or CompuGroup Medical in this case, manages the server hardware, virtualization layers, storage, and networks.

A three-tenant architecture

The RaDiCo IS architecture is distributed over 3 tenants (Fig. 3).

- *The local tenant* corresponds to the on-premises infrastructure located within the Trousseau hospital, Paris, France. It hosts the non-production environments for development, integration and testing of RaDiCo IS applications.
- *The public cloud tenant* corresponds to the infrastructure hosted and managed by the Microsoft Azure provider. It hosts all the applications that do not have any patient-related information.
- *The HDS tenant* corresponds to a private Health Data Hosting (HDS) environment approved by the French Digital Agency for Health (https//esante. gouv.fr), managed by the CompuGroup Medical (CGM) company. It hosts all the applications that contain patient-related data (personal or medical data).

RaDiCo IS applications

The RaDiCo IS is made up of a set of seven applications appearing in the capability map, covering the required functionalities that are distributed across the different tenants. These applications are as follows.

Portal and Azure Active Directory (AD)

The RaDiCo portal is a « software as a service» application provided by Microsoft on the Azure platform. It authenticates users and adapt the display of the applications that users can access to. Access rights are configured on the RaDiCo AD, and synchronised with the Azure AD.

PIST Dataware	Back Offi	ice 🔵 I Data Ar	REDCap Ad	ctiv	ve Directory	С	EGCS	Medica	ations	
Coho	Follow-up and reporting			Biometry and statistical analysis						
e-CRF r	Vérification of user's rights				Data acquisition					
Data entry form Format settings Management and validity checks			Extraction of the users attached to each cohort			Da	ta freezing	Data ext	traction	Data loading and storage
e-Cohorts data entry			Monitoring of cohorts progression		Data management and treatment					
Data entry in the Data validation e-CRF		ation	Inclusions tracking	Monitoring the data progress & completeness		Rel cor qu	liability ntrol and ality upgrade	Formats structur transfor	and es mations	Data aggregation
Data	correction		Monitoring of	м	anagement		Data	a analysis and reports		
Data qu	uality control		schedules	of	data ailability	Sta	Statistical analysis		Statistical reports	
Control of data Completeness	Managemen data consist checks	nt of ency		availabili		nty			indicators	
Queries for data clarification or correction						analyses				
Patients Users				Data repository		Access rights and security				
Patients' management Users m		anagement Management		of drugs Accou		Accourt	unt directory			
Administration Patients' Administr of patients' search repositor repository		ration of the users' Y V		Synchronisation with the WHO repository	n Drugs'search		Administration of the users' accounts and identifiers			
Management of Collection pseudonyms informat institution		n of professionals' ion (speciality, ms,) Management repository		of the cohorts' Adminis		stration of the groups				
Generation of C pseudonyms c F	Generation of Correspon- pseudonyms dence to Inclusion patients' pseudonyms		nt to a cohort groups and roles		Aministration o the cohorts' repository	f Administration of the inclusion groups repository		Settings of the access rights to the applications		
Management of consents Manage authoris		ment of ations		Management of the		Authentication				
Creation of I consent of forms	Declaration of consent	Administ (CRTs, CR	Administration of the roles (CRTs, CRAs, CRM,)		Administration of the institutions' repository			Users' authentications		
Patients'recruitment Manager (users an		agement of the Care-Team rs and patients)					Single S	ign-on		
Validation of patients' Manager eligibility authorise the patie		ment of the sations of the roles on ients					Strong a	authenticat	tion	
Fig. 2 RaDiCo Capa EGCS, Medicat	ability map and tions, Data	the corresp warehouse	oonding RaDiCo appl e, OData Analytic	licati s too	ions. PIST, Ols		Back Office,	REDCap	, Acti	ve Directory, 🔵



Fig. 3 RaDiCo SI, a three tenant architecture. (Development: DEV; Tests: TEST; Integration: INT; Pre-production: PREPROD; Production: PROD)

• Back-office (BO)

The BO application includes repositories for RaDiCo IS users, cohorts, inclusion groups, institutions, as well as their specific roles. It does not contain medical or nominative data. The BO manages user accounts by controlling the users' functional authorizations with respect to the cohorts and inclusion groups they are concerned with, according to a rights matrix. An inclusion group links patients and users, regardless of their geographical place of practice. The BO also includes a rights matrix related to medical data in REDCap (see below).

• The Patient Identifier System and Translation (PIST)

The PIST application, designed by RaDiCo, is the patients' repository. It stores three important pieces of information: Patient identifiers, the Global Unique Identifier (GUID), and the unique RD identifier (IdMR) [6], the patient's RD pseudonym, which gives access to medical data, patients' personal data and to the care-team. A care-team lists all healthcare professionals authorised to access data of patients they care. PIST does not host medical data. This application is built on a "Security by design" logic, separating personal data from medical information stored in

REDCap. Stored data are encrypted for and within the database.

• Electronic Gateway Controller Service (EGCS)

The EGCS orchestrates the patient creation/search processes between the different applications. It links the GUID code, the RD identifier and the RaDiCo code (see below). These lists are hosted within the EGCS in the form of a cross reference between the codes of each patient and the users according to their specific rights.

• Thériaque

It is the repository of medicines available in France, synchronised with the International Non-proprietary Names of the WHO [18]. It standardises medical treatments coded in REDCap.

• REDCap

It is a secure, web-based application [14] from the Vanderbilt University, USA. It has been designed to support data capture for research studies. REDCap contains all the medical data of patients included and followed in the RaDiCo cohorts. A RaDiCo code for each patient is associated with the medical data. Access to PIST and EGCS is possible via an extension module called "hook" that allows the execution of a



code or a program. Audit trails are available to track data manipulations. Automated export procedures for seamless data downloads are possible towards common statistical packages. Procedures for importing data from external sources are also available. REDCap does not host nominative data.

• Datawarehouse (DWH)

The RaDiCo DWH is a relational database hosted in the private cloud. It collects data from the cohorts with the primary objective of data analysis. Like all DWHs, it is defined as a set of subject-oriented, integrated, time-varying and non-volatile data. Stored data are not nominative.

Data architecture

The macroscopic data model of the RaDiCo IS, associated with its applications, can be broken down according to its functional entities:

- e-cohorts for clinical studies.
- Professionals' roles on the RaDiCo platform.
- Institutions with users place of practice.
- Inclusion groups that link patients and users, regardless of their geographical place of practice.

- Users of RaDiCo IS.
- Groups characterised by the authorised groups configured in the AD.
- Care-teams: all users authorised to view or intervene on patients' data.
- Patients' nominative data.
- Patients' medical data.
- Medications.

The general data architecture is modelled as depicted in Fig. 4. Most importantly, the security and anonymity of patient information are ensured by segregation between a patient's nominative data (hosted in PIST only) and their medical data (hosted in REDCap only), thereby defining a security by design. The overall architecture of the RaDiCo IS appears on Fig. 5.

Security

The security of the RaDiCo IS is based on a certified HDS, a strong and multi-factor authentication system, the fine-tuned management of access rights and digital signature mechanisms for data leaving the IS. Security mechanisms provide access control to all environments through the Azure portal application. A second security is added for access to the production environments of the HDS environment via the multi-factor authentication.



Fig. 5 RaDiCo Information System. (Cloud Microsoft AZURE; CHDH: Certified Health Data Host; CGM: CompuGroup Medical; Portal: Users authentication & selected access to the cohorts; Back Office: Management of access rights; PIST: Patient Identification System Translator; EGCS: Electronic data capture Gateway Controller Service; REDCap: Research Electronic Data Capture; DWH: RaDiCo Data warehouse; Thériaque: Drugs databank)

Personal and medical data are hosted in the private HDS tenant. Access to patient data is authorised according to a double correspondence: it is on the one hand the authorisation of access to patients modelled in the care-teams, and on the other hand the qualification of the user's effective rights, modelled in the rights matrix at the BO and REDCap levels.

Maintenance in operational condition (MCO)

Backups are carried out on a regular basis. The number of environments is suitable for software development, integration, staging and production. The GitHub tool is used to manage software configuration and code versions.

The RaDiCo cohorts

Inclusion of prevalent and incident cases started at the end of 2016 after all scientific, legal and regulatory formalities had been completed. This preparation stage was particularly long due to the French cumbersome procedures, which since then have been simplified. As of March 2024, 8063 patients have been included within the 13 RD e-cohorts, covering 67 RDs (Table 3). The principal investigators (PIs) of these cohorts belong to 10 RD European Reference Networks [19]. They contributed to the European Joint Program on RDs (EJP-RD) [20], which recently evolved into the ERDERA [21] partnership (European Rare Diseases Research Alliance). Many original results have been obtained in relation with the secondary objectives of the RaDiCo cohorts (Table 4). The data in relation with the primary objectives are currently under analysis.

Discussion

The RaDiCo IS was designed to support RD e-cohorts with innovative and general principles including factorization, pooling of resources, tools, skills and know-how, industrialization of processes, as well as quality control and security. Several implementation criteria were retained: quality, security, scalability, sustainability, and cost-efficiency.

The goal of the platform is dual. On one side, it offers a technical and IS support to maintain the running cohorts and answer to the questions defined by the research protocol. On the other side it enables running new cohorts projects. The technical support brought by the platform and its IS are structures dedicated to last within the institutional framework of Inserm. The infrastructure may be shared and pooled at the European or international level according to needs.

We have set up an e-Health & IS team to develop modern means of constructing e-cohorts with security controls supported by professionals in the field. Software and technical developments were carried out with service providers. As detailed above, the architecture includes

Cohorts' acronyms	Diseases	RDHN	ERN
AC-OEIL	Congenital Eye defects	SENSGENE AnDDI-Rares	ERN-Eye
AcoStill	Still's disease	FAI2R	ERN-Rita
COLPAC	Low Phospholipid-Associated Cholelithiasis syndrome	FILFOIE	
DCP	Primary Ciliary Dyskinesia	RESPIFIL	ERN-Lung
ECYSCO	Cystinosis	ORKID	ERKnet
EURBIO-Alport	Alport Syndrome	ORKID	ERKnet
FARD	Rare Skin Disease Burden	FIMARAD	ERN Skin
GenIDA	Intellectual deficiency and autism spectrum disorders	DéfiScience	
IDMet	Imprinting Disorders	FIRENDO OSCAR DéfiScience	ENDOERN
MPS	Mucopolysaccharidoses	G2M	MetabERN
PID	Idiopathic Interstitial Pneumonia	RESPIFIL	ERN-Lung
PP	Periodic Paralysis	FILNEMUS	Euro-NMD
SEDVasc	Vascular Ehlers-Danlos	FAVA-Multi	VASCERN

Table 3 The 13 RaDiCo cohorts by alphabetic order of their acronyms, with the names of their Rare Disease Healthcare Network (RDHN) and their contribution to RDs European Research Networks (ERNs)

several original modules designed by RaDiCo: PIST, Back Office and EGCS, all upstream of REDCap.

Other significant advances include the use of the unique and non-reversible RD identifier we have developed, the use of webservices, the possibility for patients to enter self-questionnaires, or to connect mobile objects.

The RaDiCo IS uses standards that favour interoperability of data and services, including the use of ontologies and FAIR principles. It complies with the European directive on the GDPR. Of note, the separation of the applications between the three tenants was chosen for cost reasons because HDS, with equal characteristics, costs around four times as much as public cloud hosting on Microsoft Azure. The data are hosted by an approved HDS (CompuGroup Medical, CGM), complying with 4 major criteria: availability, confidentiality, integrity, and traceability of the data. An original feature of the IS is linked to its IaaS built in the cloud. The security and anonymity of patient information are segregated between patient's nominative data (hosted in PIST only) and their medical data (hosted in REDCap only), thus defining a security by design.

A datawarehouse was created downstream of REDCap to collect reorganised data from the production database. This DWH is used for data extractions and analyses. It enables collecting data from the French National Health Data Hub and exchanges data with the ERNs.

The RaDiCo IS is trans-institutional. It currently covers 180 RD reference centres, regardless of their location. Most importantly, it has been designed to be able to integrate many more cohorts and RD reference centres without modifying its IS. Moreover, it is also available for including RD registries and their patients' follow-up. The benefit of someone joining the platform is multiple. She/he benefits from the membership of the RD community that sustains the cohort, has access to the platform resources provided by the epidemiological team and shares the IS, participates to the analysis and discussions, is included in the scientific papers, is involved in ancillary studies, and may suggest complementary investigations.

In Europe, several scattered registries are available for different RDs, but they are siloed and fragmented, and often driven by industrial purposes rather than epidemiological objectives. A European platform (EU RD platform) has been set up to standardize data collection and data exchange [81]. It is linked to the European RD registry Infrastructure (ERDRI) which supports existing registries, and new registries as well, and renders RD registries data searchable and findable [82]. The EU RD Platform also offers a data repository, the Joint Research Centre EUROCAT Central Registry on congenital anomalies [83].

National information systems for RD cohorts with available data architecture are seldom. The CEMARA IS was proposed for the French RD centres since 2007 [84]. A review of RD registries gave no specific information on the architecture of a RD IS [85]. A more recent study has formulated recommendations and guidelines for improving the quality of RD registries including guidelines for information technology infrastructure [86]. In parallel, the Euro-NMD Registry Hub has proposed a new initiative to overcome data silos and empower patients with neuromuscular diseases [87]. It is dedicated to connecting several platforms for neuromuscular diseases linked through FAIR data points.

According to recent studies, costs related to RDs are very high [88]. In this respect, the sustainability of

Table 4 International publications so far associated with the RaDiCo cohorts classified by main topics

Discovery of new disease genes	TTC12 loss-of function mutations cause primary ciliary dyskinesia and unveil distinct dynein assembly in motile cilia vs. flagella [22]
	clearance [23] Mutations in outer dynein arm heavy chain DNAH9 cause motile cilia defects and situs Inversus
	[24] Mutations in DNAJB13, encoding an HSP40 family member, cause primary ciliary dyskinesia
	and male infertility [25]
	ination and proteosomal degradation [26]
	Bi-allelic variants in WNT7B disrupt the development of multiple organs in humans [27] First evidence of SOX2 mutations in Peter's anomaly: lessons from molecular screening of 95
	Evaluation of somatic and/or germline mosaicism in congenital malformation of the eye [29]
	Cause Isolated Male Infertility Due to Asthenozoospermia [30]
Disease management treatment assessment and prognosis	Vascular Ehlers-Danlos syndrome – Long-term observational study [31] An international cohort study spanning five decades assessed outcomes of perphropatic Cystino-
	sis [32]
	Expectations about treatment of idiopathic pulmonary fibrosis: comparative survey of patients, carers, and physicians (the RESPIRATORY French study) [33]
	Follow-up and management of chronic rhinosinusitis in adults with primary ciliary dyskinesia: review and experience of our reference centers [34]
	Growth Patterns and Outcomes of Growth Hormone Therapy in Patients with Acrodysostosis [35] Height and Body Mass Index in Molecularly Confirmed Silver-Russell Syndrome and the Long-
	Impact of a Rare Respiratory Diseases Reference Centre Set-up on Primary Ciliary Dyskinesia Care
	Pathway [37]
	Obeticholic Acid as a Second-Line Treatment for Low Phospholipid-Associated Cholelithiasis Syndrome [38]
	Prevalence and Course of Disease after Lung Resection in Primary Ciliary Dyskinesia: A Cohort & Nested Case–Control Study [39]
	Recombinant Growth Hormone Improves Growth and Adult Height in Patients with Maternal Inactivating GNAS Mutations [40]
	Spontaneous Cervical Artery Dissection in Vascular Ehlers-Danlos Syndrome: A Cohort Study [41] The Impact of Lockdown on Young People with Genetic Neurodevelopmental Disabilities: A Study with the International Participatory Database GenIDA [42] Treatment of Idiopathic Pulmonary Fibrosis with Capsule or Tablet Formulations of Pirfenidone in the Pagel Life French Page [42]
Pathophysiology, diagnostic and clinical approaches	Accuracy of clinical diagnostic criteria for patients with vascular Ehlers-Danlos syndrome in a ter-
	tiary referral centre [44] Functional assessment and phenotypic heterogeneity of SFTPA1 and SFTPA2 mutations in inter- stitial lung diseases and lung cancer [45]
	Pulmonary fibrosis in children [46]
	Chronic interstitial lung diseases in children: diagnosis approaches [4/] Pulmonary hemosiderosis in children with Down syndrome: a national experience [48]
	Genetic causes and clinical management of pediatric interstitial lung diseases [50] Assessment of arterial damage in vascular Ehlers-Danlos syndrome: a retrospective multi centre
	cohort [51]
	Clinical and functional heterogeneity associated with the disruption of retinoic receptor beta [52] Clinical, genetic and biochemical signatures of RBP4-related ocular malformations [53]
	in the RaDiCo-ILD Cohort [54]
	bidities [55]
	Neonatal and Early Infancy Features of Patients With Inactivating PTH/PTHrP Signaling Disorders/ Pseudohypoparathyroidism [56]
	Otological Manifestations in Adults with Primary Ciliary Dyskinesia: A Controlled Radio-Clinical Study [57]
	Standardised Clinical Data from Patients with Primary Ciliary Dyskinesia: FOLLOW-PCD [58] Incontinentia pigmenti burden scale [59]
	Burden of albinism: development and validation of a burden assessment tool [60] Burden of adult neurofibromatosis 1: burden assessment tool [61]
	Health-related quality of life in infants and children with interstitial lung disease [62]

Table 4 (continued)

Genotype-phenotype relationship	Infertility in an adult cohort with primary ciliary dyskinesia: phenotype-gene association [63] Primary ciliary dyskinesia gene contribution in Tunisia: Identification of a major Mediterranean allele [64]
	Alport syndrome: a unified classification of genetic disorders of collagen IV @345 [65] Genetics of anophthalmia and microphthalmia. Part 1: Non-syndromic anophthalmia/microph- thalmia [66]
	Deep phenotyping, including quantitative ciliary beating parameters, and extensive genotyping in primary ciliary dyskinesia [67]
	Evaluation of somatic and/or germ line mosaicism in congenital malformation of the eye [68] Frequency of de novo variants and parental mosaicism in vascular Ehlers-Danlos syndrome [69] High Nasal Nitric Oxide, Cilia Analyses, and Genotypes in a Retrospective Cohort of Children with Primary Ciliary Dyskinesia [70] Motile Cilia and Airway Disease [71] Topological data analysis reveals genotype, phonotype relationships in primary ciliary dyskinesia
	[72]
Methodological aspects	Federating patients identities: the case of rare diseases [6] Cerberus, an access control scheme for enforcing least privilege in patient cohort study plat- forms [73]
	National registries of rare diseases in Europe: an overview of current situation and experiences [74]
	Recommendations for improving the quality of rare disease registries [75] Data quality in rare diseases registries [76]
	GenIDA, a Participatory Patient Registry for Genetic Forms of Intellectual Disability Provides Detailed Caregiver-Reported Information on 237 Individuals with Koolen-de Vries Syndrome [77] ConIDA: An International Participatory Database to Cain Knowledge on Health Issues Palated
	to Genetic Forms of Neurodevelopmental Disorders [78]
	Registries and Collaborative Studies for Primary Ciliary Dyskinesia in Europe [79] The Disease-Specific Clinical Trial Network for Primary Ciliary Dyskinesia: PCD-CTN [80]

e-cohort platforms should be encouraged to improve RD research and care. Over 50% of the RaDiCo platform is financed by academic or industrial contracts, but institutional funding remains essential to its sustainability.

Conclusion

RaDiCo provides an innovative platform for RDs, enabling the parallel management of 13 e-cohorts that currently include more than 8000 RD patients. The IS is based on a three-tenant architecture developed according to a cloud computing Infrastructure as a Service. Several modules have been created to ensure data quality and security. Interoperability and FAIR principles have been privileged. The platform can include new national or European RD cohorts, without modifying its architecture.

Abbreviations

Active Directory Back Office
(Banque nationale de données maladies rares) National Rare Disease Data Bank
Clinical Data Interchange Standards Consortium
CompuGroup Medical Company
Datawarehouse
Electronic Gateway Controller Service
European Joint Program on RDs
European Rare Diseases Research Alliance
European Reference Networks
Findability, Accessibility, Interoperability, Reusability
Global Unique Identifier

HDS	Health Data Hosting
IdMR	(Identifiant Maladie Rare) unique RD identifier
laaS	Infrastructure as a Service
KPI	Key Performance Indicator
PI	Principal investigator
PIST	Patient Identification System and Translation
R	Language and environment for statistical computing and graphics
RACI	Responsible, Accountable, Consulted, Informed
RDHN	Rare Disease Healthcare Networks
RDs	Rare diseases
REDCap	Research Electronic Data Capture
SAS	Statistical analytical software
SOA	Service-oriented architecture
SOP	Standardized operating procedures
XML	Extensible Markup Language

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13023-025-03629-z.

Supplementary Material 1.

Acknowledgements

Inserm is the sponsor of the RaDiCo cohorts. Study data were collected and managed using REDCap (Research electronic data capture) hosted at the Inserm/Sorbonne Université research unit UMR_S933 RaDiCo. REDCap is a secure, web-based application from the Vanderbilt University, USA, designed to support data capture for research studies, providing an intuitive interface for validated data entry; audit trails for tracking data manipulation and export procedures; automated export procedures for seamless data downloads to common statistical packages; and procedures for importing data from external sources. We warmly acknowledge the RaDiCo team who contributed to the implementation and follow-up of the platform. The Pls of the RaDiCo cohorts are actors of the RaDiCo Program, namely: Christine Bodemer, Patrick

Calvas, Nicolas Chassaing, Annick Clement, Christian Corpechot, Vincent Cottin, Estelle Escudier, Bruno Fautrel, Sophie Georgin Lavialle, Laurence Heidet, Bénédicte Héron, Xavier Jeunemaître, Bertrand Knebelmann, Agnès Linglart, Bernard Maitre, Jean-Louis Mandel, Irène Netchine, Aude Servais, and Savine Vicart, as well as their colleagues and teams, who contribute to the RaDiCo program.

Author contributions

Conceptualization, S.A, P.L., S.G.; methodology, P.L., S.A, SG; validation, S.A, P.L., S.G. A.C.; resources, S.A., P.L., A.C., S.G.; data curation, S.G.; writing original draft preparation, P.L., S.A., S.G.; review and editing, S.A, P.L., S.G., A.C.; supervision, P.L., S.A.; project administration, S.A., P.L., S.G., A.C.; funding acquisition, S.A., P.L., S.G., A.C. All authors have read and agreed to the published version of the manuscript.

Funding

The current funding is provided by the National Research Agency under the specific programme « Investment for the Future», Cohort grant agreement ANR-10-COHO-0003. It is coupled with complementary fundings mainly coming from the pharmaceutical industry. The financial contribution of both Research National Agency and pharmaceutical industry is administered by INSERM. The funds finance the platform infrastructure and its members, technicians, project managers, clinical research assistants and technicians, datamanagers, methodologists, and computer scientists.

Availability of data and materials

The data of the cohorts are not available presently since they are still running, and primary objectives have not been analyzed yet. The datasets analyzed will be available from the PIs of the corresponding cohorts on reasonable request.

Declarations

Ethics approval and consent to participate were given for each of the 13 RaDiCo cohorts

Ethics approval: an agreement has been requested for each RaDiCo cohort. Concerning the consent, the need for approval was waived. The ethics committee that approved each study is given in the table below. The committee's reference number was delivered for each cohort as presented below. Moreover, we obtained the agreement of the Commission Nationale Informatique et Libertés (CNIL) and from the Comité consultatif pour le traitement des données en Santé. The agreement was obtained for each cohort as stated below.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

¹Childhood Genetic Diseases Laboratory, Université de Montpellier, Inserm U933, 26 rue Arnold Netter, 75012 Montpellier, France. ²Chilhood Genetic Diseases Laboratory, Hôpital A. Trousseau, Inserm, Paris, France. ³Chilhood Genetic Diseases Laboratory, Department of Paediatric Respiratory Medicine, Hôpital A. Trousseau, Assistance Publique Hôpitaux de Paris, Sorbonne Université, Inserm, Paris, France. ⁴Chilhood Genetic Diseases Laboratory, Hôpital A. Trousseau, Sorbonne Université, Inserm, Paris, France.

Received: 19 July 2024 Accepted: 19 February 2025 Published online: 08 April 2025

References

- Nguengang Wakap S, Lambert DM, Olry A, et al. Estimating cumulative point prevalence of rare diseases: analysis of the Orphanet database. Eur J Hum Genet. 2020;28:165–73.
- 2. https://www.omim.org/statistics/geneMap. Accessed 01 May 2024.
- https://sante.gouv.fr/soins-et-maladies/prises-en-charge-specialisees/ maladies-rares/article/l-offre-de-soins

- https://sante.gouv.fr/IMG/pdf/1_P_LANDAIS_R_CHOQUET_BaMaRa-BNDMR_19mars2013.pdf. Accessed 01 May 2024.
- https://www.bndmr.fr/publication-du-set-de-donnees-minimal-malad ies-rares-national/. Accessed 01 May 2024.
- Maaroufi M, Landais P, Messiaen C, et al. Federating patients identities: the case of rare diseases. Orphanet J Rare Dis. 2018;13(1):199.
- Choquet R, Maaroufi M, de Carrara A, et al. A methodology for a minimum data set for rare diseases to support national centers of excellence for healthcare and research. J Am Med Inform Assoc. 2015;22(1):76–85.
- Amselem S, Gueguen S, Weinbach J, et al. RaDiCo, the French national research program on rare disease cohorts. Orphanet J Rare Dis. 2021;16(1):454.
- 9. https://www.orpha.net. Accessed 01 May 2024.
- 10. https://www.meddra.org. Accessed 01 May 2024.
- 11. https://hpo.jax.org/app/. Accessed 01 May 2024.
- 12. https://www.orphadata.com/ordo/. Accessed 01 May 2024.
- 13. https://www.cdisc.org/. Accessed 01 May 2024.
- Harris PA, Taylor R, Minor BL, et al. REDCap consortium, the REDCap consortium: building an international community of software partners. J Biomed Inform. 2019;95: 103208. https://doi.org/10.1016/j.jbi.2019. 103208.
- 15. https://www.nature.com/articles/sdata201618. Accessed 01 May 2024.
- 16. https://www.equator-network.org/reporting-guidelines/strobe/. Accessed 01 May 2024.
- 17. http://www.sciencedirect.com/science/article/pii/S1532046408001226. Accessed 01 May 2024.
- https://www.who.int/teams/health-product-and-policy-standards/inn. Accessed 01 May 2024.
- https://health.ec.europa.eu/european-reference-networks/rare-diseases_ en#european-reference-networks. Accessed 01 May 2024.
- 20. https://www.ejprarediseases.org. Accessed 01 May 2024.
- 21. https://www.ejprarediseases.org/erdera/. Accessed 01 May 2024.
- Thomas L, Bouhouche K, Whitfield M, et al. TTC12 loss-of-function mutations cause primary ciliary dyskinesia and unveil distinct dynein assembly mechanisms in motile cilia versus flagella. Am J Hum Genet. 2020;106(2):153–69. https://doi.org/10.1016/j.ajhg.2019.12.010.[PMCfr eearticle][PubMed][CrossRef][GoogleScholar].
- Bustamante-Marin XM, Yin WN, Sears PR, et al. Lack of GAS2L2 causes PCD by impairing cilia orientation and mucociliary clearance. Am J Hum Genet. 2019;104(2):229–45. https://doi.org/10.1016/j.ajhg.2018.12.009. [PMCfreearticle][PubMed][CrossRef][GoogleScholar].
- 24. Fassad MR, Shoemark A, Legendre M, et al. Mutations in outer dynein arm heavy chain DNAH9 cause motile cilia defects and situs inversus. Am J Hum Genet. 2018;103(6):984–94. https://doi.org/10.1016/j.ajhg.2018.10. 016.[PMCfreearticle][PubMed][CrossRef][GoogleScholar].
- El Khouri E, Thomas L, Jeanson L, et al. Mutations in DNAJB13, encoding an HSP40 family member, cause primary ciliary dyskinesia and male infertility. Am J Hum Genet. 2016;99(2):489–500. https://doi.org/10.1016/j. ajhg.2016.06.022.[PMCfreearticle][PubMed][CrossRef][GoogleScholar].
- Holt RJ, Young RM, Crespo B, et al. De novo missense variants in FBXW11 cause diverse developmental phenotypes including brain, eye, and digit anomalies. Am J Hum Genet. 2019;105(3):640–57. https://doi.org/10. 1016/j.ajhg.2019.07.005.[PMCfreearticle][PubMed][CrossRef][GoogleScho lar].
- Bouasker S, Patel N, Greenlees R, et al. Bi-allelic variants in WNT7B disrupt the development of multiple organs in humans. J Med Genet. 2023;60:294–300.
- Chesneau B, Aubert-Mecca M, Fremont F, et al. First evidence of SOX2 mutations in Peter's anomaly: lessons from molecular screening of 95 patients. Clin Genet. 2022;101:494–506.
- 29. Chesneau B, Ivashchenko V, Habib C, et al. Evaluation of somatic and/or germline mosaicism in congenital malformation of the eye. Eur J Human Genet. 2023;31:526–30. https://doi.org/10.1038/s41311-022-01234-3.
- Whitfield M, Thomas L, Bequignon E, et al. Mutations in DNAH17, encoding a sperm-specific axonemal outer dynein arm Heavy Chain, cause isolated male infertility due to asthenozoospermia. Am J Human Genet. 2019;105:198–212. https://doi.org/10.1016/j.ajhq.2019.04.015.
- Frank M, Adham S, Seigle S, et al. Vascular Ehlers-Danlos syndrome: longterm observational study. J Am Coll Cardiol. 2019;73(15):1948–57. https:// doi.org/10.1016/j.jacc.2019.01.058.[PubMed][CrossRef][GoogleScholar].

- Emma F, van't Hoff W, Hohenfellner K, et al. An international cohort study spanning five decades assessed outcomes of nephropatic Cystinosis. Kidney Int. 2021;100(5):1112–23. https://doi.org/10.1016/j.kint.2021.06. 019.
- Cottin V, Bergot E, Bourdin A, et al. Expectations about treatment of idiopathic pulmonary fibrosis: comparative survey of patients, carers, and physicians (the RESPIRATORY French study). Respir Med Res. 2021;79: 100811. https://doi.org/10.1016/jresmer.2020.100811.
- Bequignon E, Dupuy L, Escabasse V, et al. Follow-up and management of chronic rhinosinusitis in adults with primary ciliary dyskinesia: review and experience of our reference centers. J Clin Med. 2019;8:1495.
- Ertl DA, Mantovani G, de Nanclares GP, et al. Growth patterns and outcomes of growth hormone therapy in patients with acrodysostosis. J Endocrinol Invest. 2023;46:1673–84. https://doi.org/10.1007/ s40618-023-02026-2.
- Lokulo-Sodipe O, Giabicani E, Canton APM, et al. Height and body mass index in molecularly confirmed silver-russell syndrome and the long-term effects of growth hormone treatment. Clin Endocrinol. 2022;97:284–92. https://doi.org/10.1111/cen.14715.
- Epaud S, Epaud R, Salaün-Penquer N, et al. Impact of a rare respiratory diseases reference centre set-up on primary ciliary dyskinesia care pathway. Eur Respir J. 2022;59:2102413. https://doi.org/10.1183/13993003. 02413-2021.
- Soret PA, Lemoinne S, Mallet M, Ben Belkacem K, Chazouillères O, Corpechot C. Obeticholic acid as a second-line treatment for low phospholipid-associated cholelithiasis syndrome. Aliment Pharmacol Ther. 2024;59:113–7. https://doi.org/10.1111/apt.17761.
- 39. On behalf of the Israeli PCD Consortium, on behalf of the Italian PCD Consortium, on behalf of the Swiss PCD Group, on behalf of the French Reference Centre for Rare Lung Diseases, on behalf of the PCD Italian Consortium, on behalf of the French Reference Centre for Rare Lung Diseases, Panayiotis Kouis, et al. Prevalence and Course of Disease after Lung Resection in Primary Ciliary Dyskinesia: A Cohort & Nested Case-Control Study. *Respir Res* 2019 ;20:212. https://doi.org/10.1186/s12931-019-1183-y.
- Ertl DA, Perez de Nanclares G, Jüppner H, et al. Recombinant growth hormone improves growth and adult height in patients with maternal inactivating GNAS mutations. Eur J Endocrinol. 2023;189:123–31. https://doi.org/10.1093/ejendo/lvad085.
- Adham S, Billon C, Legrand A, et al. Spontaneous cervical artery dissection in vascular Ehlers-Danlos syndrome: a cohort study. Stroke. 2021;52:1628–35. https://doi.org/10.1161/STROKEAHA.120.032106.
- 42. Coutelle R, Boedec M, Vermeulen K, et al. The impact of lockdown on young people with genetic neurodevelopmental disabilities: a study with the international participatory database GenIDA. BMC Psychiatry. 2022;22:572. https://doi.org/10.1186/s12888-022-04213-6.
- Cottin V, Guéguen S, Nunes H, et al. Treatment of idiopathic pulmonary fibrosis with capsule or tablet formulations of pirfenidone in the reallife French RaDiCo-ILD cohort. Adv Ther. 2021;39:405–20. https://doi. org/10.1007/s12325-021-01961-x.
- Henneton P, Albuisson J, Adham S, et al. Accuracy of clinical diagnostic criteria for patients with vascular ehlers-danlos syndrome in a tertiary referral centre. Circ Genom Precis Med. 2019;12(3):e001996.
- Legendre M, Butt A, Borie R, et al. Functional assessment and phenotypic heterogeneity of SFTPA1 and SFTPA2 mutations in interstitial lung diseases and lung cancer. Eur Respir J. 2020. https://doi.org/10.1183/ 13993003.02806-2020.[PubMed][CrossRef][GoogleScholar].
- Nathan N, Sileo C, Thouvenin G, et al. Pulmonary fibrosis in children. J Clin Med. 2019;8(9):1312. https://doi.org/10.3390/jcm8091312.[PMCfr eearticle][PubMed][CrossRef][GoogleScholar].
- Nathan N, Berdah L, Borensztajn K, Clement A. Chronic interstitial lung diseases in children: diagnosis approaches. Expert Rev Respir Med. 2018;12:1051–60.
- Alimi A, Taytard J, Abou Taam R, et al. Pulmonary hemosiderosis in children with down syndrome: a national experience. Orphanet J Rare Dis. 2018;13(1):60. https://doi.org/10.1186/s13023-018-0806-6.[PMCfr eearticle][PubMed][CrossRef][GoogleScholar].
- Nathan N, Sileo C, Calender A, et al. French sarcoidosis group (GSF); silicosis research group. Paediatric sarcoidosis Paediatr Respir Rev. 2019;29:53–9. https://doi.org/10.1016/j.prrv.2018.05.003.[PubMed] [CrossRef][GoogleScholar].

- Nathan N, Borensztajn K, Clement A. Genetic causes and clinical management of pediatric interstitial lung diseases. Curr Opin Pulm Med. 2018;24(3):253–9. https://doi.org/10.1097/MCP.000000000000471. [PubMed][CrossRef][GoogleScholar].
- Adham S, Legrand A, Bruno RM, et al. Assessment of arterial damage in vascular Ehlers-Danlos syndrome: a retrospective multi centre cohort. Front Cardiovascular Med. 2022;9:953894. https://doi.org/10.3389/ fcvm.2022.953894.
- Caron V, Chassaing N, Ragge N, et al. Clinical and functional heterogeneity associated with the disruption of retinoic receptor beta. Genet Med. 2023;25: 100856.
- Plaisancié J, Martinovic J, Chesneau B, et al. Clinical, genetic and biochemical signatures of RBP4-related ocular malformations. J Med Genet. 2023;61:84–92.
- Cottin V, Gueguen S, Jouneau S, et al. Impact of gender on the characteristics of patients with idiopathic pulmonary fibrosis included in the RaDiCo-ILD cohort. Respiration. 2021. https://doi.org/10.1159/00051 8008.
- Dong C, Condat B, Picon-Coste M, et al. Low-phospholipid-associated cholelithiasis syndrome: prevalence, clinical features, and comorbidities. JHEP Rep. 2021;3: 100201. https://doi.org/10.1016/j.jhepr.2020.100201.
- Del Sindaco G, Berkenou J, Pagnano A, et al. Neonatal and early infancy features of patients with inactivating PTH/PTHrP signaling disorders/ pseudohypoparathyroidism. J Clin Endocrinol Metab. 2023;108:2961–9. https://doi.org/10.1210/clinem/dgad236.
- Alexandru M, de Boissieu P, Benoudiba F, et al. Otological manifestations in adults with primary ciliary dyskinesia: a controlled radio-clinical study. J Clin Med. 2022;11:5163. https://doi.org/10.3390/jcm11175163.
- Goutaki M, Papon JF, Boon M, et al. Standardised clinical data from patients with primary ciliary dyskinesia: follow-PCD. ERJ Open Res. 2020;6:00237–2019. https://doi.org/10.1183/23120541.00237-2019.
- Taieb C, Hadj-Rabia S, Monnet J, et al. Incontinentia pigmenti burden scale: designing a family burden questionnaire. Orphanet J Rare Dis. 2019;14:271. https://doi.org/10.1186/s13023-019-1234-y.
- Morice-Picard F, Taïeb C, Marti A, et al. Burden of albinism: development and validation of a burden assessment tool. Orphanet J Rare Dis. 2018;13(1):162.
- Armand ML, Taieb C, Bourgeois A, et al. Burden of adult neurofibromatosis 1: development and validation of a burden assessment tool. Orphanet J Rare Dis. 2019;14(1):94. https://doi.org/10.1186/s13023-019-1067-8.
- Lauby C, Boelle PY, Abou Taam R, et al. Health-related quality of life in infants and children with interstitial lung disease. Pediatr Pulmonol. 2019;54(6):828–36. https://doi.org/10.1002/ppul.24308.
- Vanaken GJ, Bassinet L, Boon M, et al. Infertility in an adult cohort with primary ciliary dyskinesia: phenotype-gene association. Eur Respir J. 2017;50(5):1700314.
- 64. Mani R, Belkacem S, Soua Z, et al. Primary ciliary dyskinesia gene contribution in Tunisia: identification of a major Mediterranean allele. Hum Mutat. 2020;41(1):115–21. https://doi.org/10.1002/humu.23905.
- Kashtan CE, Ding J, Garosi G, et al. Alport syndrome: a unified classification of genetic disorders of collagen IV α345: a position paper of the Alport syndrome classification working group. Kidney Int. 2018;93:1045– 51. https://doi.org/10.1016/j.kint.2017.12.018.
- Plaisancié J, Ceroni F, Holt R, et al. Genetics of anophthalmia and microphthalmia. Part 1: non-syndromic anophthalmia/microphthalmia. Hum Genet. 2019;138(8–9):799–830.
- Blanchon S, Legendre M, Bottier M, et al. Deep phenotyping, including quantitative ciliary beating parameters, and extensive genotyping in primary ciliary dyskinesia. J med Genet. 2020;57:237–44. https://doi.org/ 10.1136/jmedgenet-2019-106424.
- Ivashchenko V, Habib C, Gaston V, et al. Evaluation of somatic and/or germ line mosaicism in congenital malformation of the eye. Eur J Hum Genet. 2023;31:526–30.
- Legrand A, Devriese M, Dupuis-Giroud S, et al. Frequency of de novo variants and parental mosaicism in vascular Ehlers-Danlos syndrome. Genet Med. 2019;21:1568–75.
- Legendre M, Thouvenin G, Taytard J, et al. High nasal nitric oxide, cilia analyses, and genotypes in a retrospective cohort of children with primary ciliary dyskinesia. Ann Am Thorac Soc. 2022;19:1704–12. https://doi. org/10.1513/AnnalsATS.202110-1175OC.

- Legendre M, Zaragosi LE, Mitchison HM. Motile cilia and airway disease. Semin Cell Dev Biol. 2020;110:19–33. https://doi.org/10.1016/j.semcdb. 2020.11.007.
- ShoemarkA RB, Legendre M, et al. Topological data analysis reveals genotype-phenotype relationships in primary ciliary dyskinesia. Eur Respir J. 2021;58(2):2002359. https://doi.org/10.1183/13993003.02359-2020.
- Parrend P, Mazzucotelli T, Colin F, et al. Cerberus, an access control scheme for enforcing least privilege in patient cohort study platforms a comprehensive access control scheme applied to the GENIDA project study of genetic forms of intellectual disabilities and autism spectrum disorders. J Med Syst. 2018. https://doi.org/10.1007/s10916-017-0844-y.
- Taruscio D, Vittozzi L, Choquet R, et al. National registries of rare diseases in Europe: an overview of current situation and experiences. Public Health Genomics. 2015;18(1):20–5. https://doi.org/10.1159/000365897.
- Kodra Y, Weinbach J, Posada-de-la-Paz M, Coi A, et al. Recommendations for improving the quality of rare disease registries. Int J Environ Res Public Health. 2018;15(8):1644.
- Kodra Y, Posada de la Paz M, Coi A, et al. Data quality in rare diseases registries. Adv Exp Med Biol. 2017;1031:149–64.
- Colin F, Burger P, Mazzucotelli T, et al. GenIDA, a participatory patient registry for genetic forms of intellectual disability provides detailed caregiver-reported information on 237 individuals with Koolen-de Vries syndrome. Gen Med Open. 2023;1: 100817. https://doi.org/10.1016/j. gimo.2023.100817.
- Burger P, Colin F, Strehle A, et al. GenIDA: an international participatory database to gain knowledge on health issues related to genetic forms of neurodevelopmental disorders. J Neural Transm. 2023;130:459–71. https://doi.org/10.1007/s00702-022-02569-3.
- Ardura-Garcia C, Goutaki M, Carr SB, et al. Registries and collaborative studies for primary ciliary dyskinesia in Europe. ERJ Open Res. 2020;6:00005–2020. https://doi.org/10.1183/23120541.00005-2020.
- Raidt J, Maitre B, Pennekamp P, et al. The disease-specific clinical trial network for primary ciliary dyskinesia: PCD-CTN. ERJ Open Res. 2022;8:00139–2022. https://doi.org/10.1183/23120541.00139-2022.
- 81. https://eu-rd-platform.jrc.ec.europa.eu/aim-of-the-platform_en. Accessed 01 May 2024.
- 82. https://eu-rd-platform.jrc.ec.europa.eu/erdri_en. Accessed 01 May 2024.
- https://eu-rd-platform.jrc.ec.europa.eu/eurocat_en. Accessed 01 May 2024.
- Landais P, Messiaen C, Rath A, et al. CEMARA task force. CEMARA an information system for rare diseases. Stud Health Technol Inform. 2010;160:481–5.
- Derayeh S, Kazemi A, Rabiei R, et al. National information system for rare diseases with an approach to data architecture: a systematic review. Intractable Rare Dis Res. 2018;7(3):156–63.
- Tarride JE, Okoh A, Aryal K, Prada C, et al. Scoping review of the recommendations and guidance for improving the quality of rare disease registries. Orphanet J Rare Dis. 2024;19(1):187. https://doi.org/10.1186/s13023-024-03193-y.
- Atalaia A, Wandrei D, Lalout N, et al. EURO-NMD registry: federated FAIR infrastructure, innovative technologies and concepts of a patientcentred registry for rare neuromuscular disorders. Orphanet J Rare Dis. 2024;19(1):66. https://doi.org/10.1186/s13023-024-03059-3.
- 88. https://everylifefoundation.org/burden-landing/. Accessed 01 May 2024.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.