MEDINFO 2019: Health and Wellbeing e-Networks for All
L. Ohno-Machado and B. Séroussi (Eds.)

doi:10.3233/SHTI190391

Establishment of a Comprehensive Information Infrastructure and a Support Organization for Rare Disease Research in Japan (RADDAR-J)

Izumi Yamaguchi[a], Yoshihiko Furusawa[b], Takahisa Kawaguchi[a], Naoko Yagishita[c], Kazumasa Tanzawa[d], Yoshihisa Yamano [c], Fumihiko Matsuda[a]

[a] Center for Genomic Medicine, Graduate School of Medicine, Kyoto University,
[b] Department of Neurology, National Center Hospital, National Center of Neurology and Psychiatry,
[c] Department of Rare Diseases Research, Institute of Medical Science, St. Marianna University School of Medicine,
[d] EPS Corporation

Abstract

There are more than 300 research groups for rare diseases in Japan. Although various clinical and genomic information of patients are being collected by the groups, the information is managed individually by each research group and the current practices for managing and sharing research data are not very efficient. Since "rare diseases" are literally rare, the understanding of the underlying disease mechanisms are incomplete and collecting a sufficient number of patients for clinical trials is difficult. Therefore, there is a need to collect and integrate the data and construct a data integration platform for rare disease research. Funded by the Japan Agency for Medical Research and Development, a national research and development project to establish a standard platform and supporting organization for rare disease registries in Japan is currently under way. In this article, we report the background, purpose, process, results, current status, and future plans of this project.

Keywords:

Rare diseases, registries, data sharing.

Introduction

In Japan, research studies on rare diseases have been conducted for over four decades, and the Ministry of Health, Labour and Welfare of Japan (hereafter, referred to as "MHLW") has been supporting these studies.

The MHLW or the Japan Agency for Medical Research and Development (hereafter called "AMED" and was established in 2015 as a funding agency for medical research in Japan) supports about 300 research groups for rare diseases in Japan. In addition, there are research groups of various sizes that are not funded by MHLW or AMED.

All of these research groups are collecting various kinds of clinical and genomic information from patients for their own purposes. However, in most cases such data are managed individually by each research group, and long-term stable management of the data or effective data sharing among research groups are not sufficiently achieved because of the burden on each research group.

Since the number of the rare disease patients is low, collecting a sufficient number of patients for clinical trials is frequently difficult. As a result, the quantity and quality of the information about rare disease patients which a single research group can collect is limited, and the understanding of the underlying disease mechanisms for rare diseases is largely unknown.

Therefore, there is a strong need for collecting and integrating the data from rare disease patients and constructing a data integration platform for rare disease research in Japan. Such a platform will contribute greatly to the advance and expansion of rare disease research.

With this in mind, a research project was launched by AMED in 2017 to establish a national research infrastructure for rare disease registries in Japan, and our research group was placed in charge of the project [1].

Methods

The purpose and the goal of this research project is to establish a national information infrastructure and management organization to enable: stable longitudinal management; effective sharing among research groups; and integration, analysis, and secondary use of the clinical and genomic information from the research participants.

The information system and the supporting organization was named "RADDAR-J" (RAre Disease DAta Registry of Japan) and the following tasks were carried out:

1. Administering a web-based survey of the rare disease research groups supported by AMED or MHLW
2. Defining the rules and regulations
3. Designing and developing the information infrastructure
4. Establishing a support system for the management of registries

Web-based survey

Modeled on a prior survey in Europe [2], a web-based survey targeting rare disease research groups supported by AMED or MHLW was conducted in order to investigate the current development and management of rare disease registries in Japan, the needs for support, and so on. The survey items were as follows:

- Characteristics of the research project (5 questions)
- Characteristics of the disease registry (25 questions)
- Information system of the registry (24 questions)
- Management and registration scheme (35 questions)
- Collected data (48 questions)
- Management of the data (18 questions)
- Biorepository (16 questions)
- Needs and expectations (2 questions)

Defining the rules and regulations

Rules and regulations necessary for collecting, integrating, sharing, and providing the clinical and genomic information from the research groups included:

- Designing the organization
- Developing the operating policy
- Developing the data-sharing policy

Designing and developing the information infrastructure

The overall design of the information infrastructure necessary for RADDAR-J was created, and the development of the system was started.

Establishing a support system for registry management

A support system for registry management was organized by:

- Establishing a secretariat of RADDAR-J
- Setting up and operating the portal site in both Japanase and English
- Preparing the standard documents for management and operation of the registry
- Holding consultation meetings for research groups about the disease registry

Results

Web-based survey

A total of 303 rare disease research groups supported by AMED or MHLW were surveyed, and 89% (271 groups) answered our web-based questionnaire, which consisted of 173 questions mentioned above. The initial analysis of the answers showed that :

- 74% (200 groups) of the responding groups had some kind of registry.
- 29% (78 groups) were operating some data entry system for registry.
- 18% (49 groups) had a problem in collecting research participants.
- 12% (32 gropus) had a problem in utilizing the collected data effectively.
- 37% (100 groups) were short-staffed (especially with regard to specialist staff).

A part of the survey data was released to the public on our portal site, and all the data obtained from the survey are now under analysis in detail.

Defining the rules and regulations

Designing the organization of RADDAR-J

Design of the organization of RADDAR-J was modeled on an international organization in Europe named "RD-Connect" (https://rd-connect.eu/), which aims to accelerate rare disease research by providing a global platform for database, registries, and biobanks [3]. Based on the result of the web-based survey, the roles of RADDAR-J were set as follows:

- Collecting, managing, and providing the information and data shared from each disease research registry
- Integrating and analyzing the data provided from the disease registries
- Supporting the management of each disease registry research
- Returning the analytic results of the integrated data by RADDAR-J to the original research groups

Considering the diversity of the data types and the importance of the patients' personal information management, it was decided that RADDAR-J should consist of three main departments (Personal Information Management department, Clinical Information Integration department, and Genomic Information Integration department) and a RADDAR-J secretariat, which controls the whole organization. A committee was formed for each department.

Developing the operating policy of RADDAR-J

Detailed operating policies of RADDAR-J, such as the organizational structure, the role of each department, the operation of each committee, the necessary specifications of the standard registry system for reseach groups, the rules for the registry system management and the data management, were developed and documented.

1. Personal Information Management department is mainly in charge of the management of patients' personal information including real names. The department performs a “name-based aggregation” of patient information within one research project group as well as between different research project groups, so that patients are not duplicated within projects and can be tracked between projects.

   Since there is no national identification number available for medical research in Japan, personal information including real names need to be used to identify each patient correctly in order to avoid double registration of a patient in one research project or to find the existence of a patient that is registered in multiple disease research projects. After the “Name-based aggregation” process, patients whose real names, dates of birth, and gender match each other will be extracted as “possibly identical patients”.
2. Clinical Information Integration department is mainly in charge of :
   - supporting research groups in the development and the management of registry.
   - collecting, integrating, and managing clinical data provided by each research group.
   - aggregating and analyzing the integrated data.
3. Genomic Information Integration department is mainly in charge of :
   - collecting and managing genomic information and omics information provided by each research group.
   - collecting some clinical information from Clinical Information Integration department necessary for genomic analysis or omics analysis, and performing an integrated analysis of these data and genomic and/or omics data.
   - supporting research groups in genomic analysis such as providing some disease analysis tools, and providing genomic and/or omics information of control groups available for disease analysis.

Developing the data-sharing policy of RADDAR-J

A data-sharing policy of RADDAR-J was developed based on “Data Sharing Policy for Facilitation of Genomic Medicine” of AMED [4]. Three forms of data-sharing are defined in the policy as follows:

1. Group sharing data

   Data will be shared only between RADDAR-J and the research group, the original source of the data.
2. Controlled-Access Data

Data will be provided to third parties based on review and approval of RADDAR-J if the provision is approved in advance by the original source of the data.

3. Open Data

   Data will be published by RADDAR-J publicly if the publication is approved in advance by the original source of the data.

In order to operate these policies, researchers should obtain written informed consent from patients or use an opt-out consent process only when the research is an academic study and opt-in procedures are impractical, as described in the national ethical guidelines.

The basic data sharing policy of RADDAR-J was decided as group sharing between each research group and RADDAR-J in order to make it easier for each reseach group to obtain agreement with patients on data sharing. Each reseach group develops their own data sharing policy with RADDAR-J and its external seconady users.

Designing and developing the information infrastructure

Based on the organizational structure and the role of each department, a system configuration of RADDAR-J was designed, in which each department holds its own dedicated system and all the systems cooperate with each other as shown in Figure 1.

- Personal Information Management System, used in the Personal Information Management department, stores and manages the patients' basic personal information including real names in encrypted form. In addition, it is able to compare the personal information without decryption for privacy protection, so that it can safely perform "name-based aggregation" of patients as necessary [5].
- In order to standardize the clinical information of each rare disease registry as much as possible, a standard registry system was prepared for each research group to use at a reasonable price, reducing the burden of the system development and management. The system is based on an electronic data capture system that one of our research collaborators had developed as a commercial product. It stores and manages various kinds of patient clinical information in de-identified form and some personal information in encrypted form, which are necessary for the operation of each registry secretariat. International standard codes such as ICD-10 and HPO (Human Phenotype Ontology) were implemented on the standard registry system to facilitate international sharing of information.
- Regarding Clinical Information Integration System that will be used in the Clinical Information Integration department, the specification including the system linkage is currently being considered, and the development will start shortly. The system collects the fixed standard clinical information from each registry system in de-identified and standardized form, and receives the result of "name-based aggregation" from the Personal Information Management System as necessary.
- Genomic Information Integration System, used in the Genomic Information Integration department, stores and manages genomic information, omics information, and some clinical information necessary for genomic analysis in de-identified form. In addition, some simple genetic analysis functions are planned to be implemented in the system. This system is currently under development. The basic functions such as storing or managing genomic information have already been developed, and additional functions such as searching for genetic mutations are under development.

Support system for the management of registries

The official website of RADDAR-J was launched in both Japanese and English (https://raddarj.org/en/). An explanation of the RADDAR-J project and the result of the web-based survey about rare disease registries in Japan are available on the site.

The RADDAR-J secretariat held an explanatory meeting about the RADDAR-J project with rare disease research groups funded by AMED or MHLW, and is conducting individual consultations for research groups that request support for developing their disease registries. For an individual consultation, members of the secretariat have an interview with the reseach group to determine their current situation and demand, and consider how to support them.

Various kinds of standard documents for developing and managing rare disease registries are being developed as shown in Table 1. They will be provided at no charge to reseach groups according to their needs.

Discussion

Background of this research project

According to AMED policy, data obtained from research and development supported by public funds is national intellectual property, and AMED places great importance to the appropriate collection, quality assurance, presevation, utilization and application of data. To that end, AMED has established "Data Sharing Policy for the Realization of Genomic Medicine" [4] and promotes data sharing. Based on the policy, AMED launched a research project for the construction of information infrastructure for rare/intractable diseases, and awarded funding to the current authors' researh group.

RADDAR-J aims to contribute to the acceleration of rare disease research by providing the infrastructure and the organization which enables effective sharing and integration of clinical and genomic information that rare disease research groups collect.

Need for support

Although the data from the web-based survey, which was conducted in advance of designing the organization and the information system of RADDAR-J, are currently under detailed analysis, the survey showed that about three quarters of the responding research groups hold some kind of registry, but that less than half of them operate some data entry system for registry. In addition, slightly more than one-third of the responding research groups are understaffed (especially lacking in specialist staff).

From these results, we concluded that support by RADDAR-J for the information system and the operation of registry would help to advance the progress of rare disease research.

What we have done so far

After the whole picture for RADDAR-J including the necessary systems and regimes were designed, the development of the information system and the supporting organization for the registry systems was prioritized. The rare disease research groups could thus smoothly launch and start using their own registries because this part is, as it were, a gate for the data into RADDAR-J.

In order to secure data provenance, a standard registry system with operational log function and user authentification function, as well as standard documents for organizational operation and procedures are offered by RADDAR-J. Using these, the research groups can manage their registries and the information appropriately and undergo monitoring when necessary.

Next, an explanatory meeting about the RADDAR-J project was held with rare disease research groups funded by AMED or MHLW, and individual consultations started to be conducted for research groups that request support for developing their disease registries. Being able to exchange frank opinions which can improve future support, these consultations were found to be good opportunities to get to know deeply the actual circumstances of the research groups.

What we are doing

Currently, the information system and the organization for integration, analysis, and secondary use of the patient data collected by the registry system is being developed. Since multiple systems of different origins and by different software manufacturers need to link to each other mainly due to budget, there are many problems to be solved such as the specification of system linkage, the specification of data integration (e.g., patient ID, project ID, data format), the system operations, and so on. Therefore, the detailed specification development has been challenging and time consuming.

There is no national identification number available for medical practice in Japan, and so harmonizing information between registries and clinical record systems is a big challenge for this project. A new system of national identification number for medical practice is recently being designed, which may enable such harmonization in Japan as in the Nordic countries [6].

It can be said that this research project is still in the midst of development and there are numerous issues that remain to be resolved. However, when complete, this infrastructure and organization will greatly fascilitate effective data sharing and promote the progress of rare disease research. International cooperation with some similar organizations in Europe is currently under cosideration.

Conclusions

Data sharing between research groups will play an important role in the field of rare disease research due to the rarity of patients. In this research project (RADDAR-J), an information infrastructure and a supporting organization are being developed for building rare disease registries and sharing patient data, which are expected to greatly contribute to the progression of rare disease research.

Table 1– Examples of the standard documents to be provided

Name of Documents
Sample of Protocol
Sample of Informed Consent Form
Sample of Informed Assent Form
Guide to modification of an already existing protocol
Guide to modification of an already existing informed consent form
Sample of information disclosure document for opt-out
Guideline for data sharing
Standard Operating Procedure for research
Operating Procedure for data management
Operating Procedure for configuration and management of EDC (Electric Data Capture)
Operating Procedure for monitoring and audit
Manual for data collecting
Manual for data entry to EDC
Database structure definition
Case Report Form
Statistical analysis plan

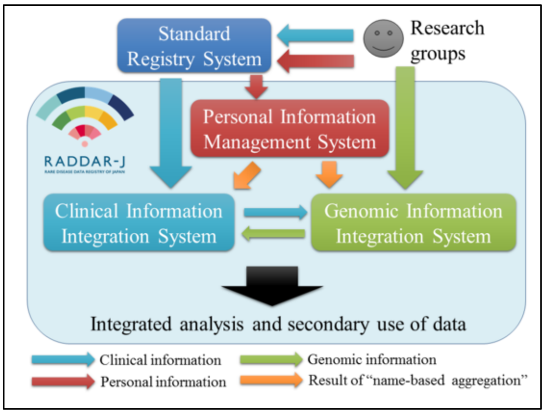

Figure 1– System configuration and data flow of RADDAR-J

Acknowledgements

RADDAR-J is a three-year research project which started in February 2017 and is supported by AMED under Grant Number JP17ek0109196 and JP18ek0109348.

References

[1] Furusawa Y, Yamaguchi I, Yagishita N, et al. National platform for Rare Diseases Data Registry of Japan. *Learn Health Systems*. 2019;e10080.

[2] Taruscio D. et al. The Current Situation and Needs of Rare Disease Registries in Europe. *Public Health Genomics* 16 (2013), 288–298.

[3] Thompson R. et al. RD-Connect: An Integrated Platform for Rare Disease Research. *Journal of General Internal Medicine*, 2014 Aug;29 Supple 3:S780-7.

[4] Japan Genomic Medicine Project Data sharing Policy for the Realization of Genomic Medicine. https://www.amed.go.jp/content/000017356.pdf

[5] Yamaguchi I. et al. Development of a real name registration and verification system for epidemiologic studies in order to avoid overlapping of patient registration. *Japan Journal of Medical Informatics*, 2017; 37(Suppl.): 939-942.

[6] Andreassen OA. eHealth provides a novel opportunity to exploit the advantages of the Nordic countries in psychiatric genetic research, building on the public health care system, biobanks, and registries. *American Journal of Medical Genetics Part B*. 2018;177B:625–629

Ethics approval

The positive ethics vote (No. R1399-2) from the ethics committee of Kyoto University Graduate School and Faculty of Medicine, Kyoto, Japan was obtained (chairman: Atsushi Asai).

Address for correspondence

Izumi Yamaguchi, M.D.
Center for Genomic Medicine,
Graduate School of Medicine, Kyoto University
Nanbusogo-Kenkyu-To-1, 5F
53 Shogoinkawahara-cho, Sakyo-ku, Kyoto 606-8507, JAPAN
Tel: (+81)-75-751-4157
FAX: (+81)-75-751-4167
E-mail: yamaguchi@genome.med.kyoto-u.ac.jp