FAQs for the Global Rare Diseases Registry Data Repository (GRDR)

1. **How was the GRDR developed, and what are its goals?**

The GRDR was initiated in 2010 as a demonstration or "proof of concept" (POC) project, and has undergone several changes in focus as lessons have been learned over the course of its development.

The GRDR program activities began following a 2010 workshop sponsored by the NIH Office of Rare Diseases Research (ORDR), titled "Advancing Rare Disease Research: The intersection of patient registries, biospecimen repositories, and clinical data".[1] This workshop explored the development of an infrastructure for an internet-based IT platform, common data elements (CDEs), and a global rare disease registry able to incorporate existing registries and establish new ones. Initially, 12 patient advocacy group-led registries were supported by the GRDR, which had the major aims of:

- Validating and implementing common data elements
- Gauging general interest from the rare diseases community in constructing registries

This POC project was completed in 2013. Some of the key points learned from the POC projects were that, given the large number of rare diseases, most of which have unmet medical needs, this approach was not scalable to more than 6,000 rare diseases. The GRDR would, therefore, need to take a broader focus to be able to inform registry building for a larger number of diseases, and that data interoperability to facilitate pooling and analyses across rare diseases would be needed.

Subsequently, and based on the lessons learned from the POC project, in 2015 NCATS issued a supplemental grant to Harvard Medical School (HMS) Department of Biomedical Informatics (DBMI) to perform a data mapping exercise for data in selected registries. The aim of this exercise was to map data contained in selected registry datasets to commonly used ontologies and vocabularies, such that data from a wide range of sources (e.g., registries, electronic health records, research datasets, etc.) could be integrated to allow for broader use and sharing of the data across rare diseases.

The mapping from the selected registries has been completed, and the data were mapped to more than 20 different ontologies, such as ICD9, ICD10, SnoMED, MeSH, among others. This exercise demonstrated that post-processing of data from registries is feasible. However, similar to lessons learned from the first phase, the second phase again underscored that this approach is not scalable, and that the GRDR would need to broaden its reach to inform registry building across the larger rare disease community.

Current goals are focused on convening the greater informatics, registry-building, and rare diseases community to work together to define best practices for registries, and to facilitate the identification and adoption of standards and practices that are able to support good quality registries able to inform therapeutics development for rare diseases.

2. **What changes are planned for the GRDR?**

First, to reflect the new focus of the GRDR, the program will undergo a name change to the "**Rare Diseases Registry Program**" (RaDaR).

[1] https://rarediseases.info.nih.gov/files/Patient_Registries_Workshop_Materials_JAN2010.pdf

Second, the upcoming phase of the registries project will focus on:

- Identifying, validating, and developing data standards, data collections and data sharing practices that can be broadly accepted and used across the rare disease registry community.
- Developing best practices and advice to build high-quality registries able to support therapeutics development, and
- Advice and standards to be made broadly available and easily accessible to the rare disease community.

Plans are currently in progress to develop a process on how to proceed to reach these broad and ambitious goals. Steps thus far include:

- The old GRDR website will be revised to the RaDaR website, and will include up-to-date information on the status of the RaDaR.
- An initial data standards meeting, titled "Data standards for registries: Facilitating the flow of information" was held on June 28, 2017. This meeting convened a group of subject matter experts from the IT, informatics and rare disease clinical areas to survey the state of rare disease registries and begin planning next steps for the RaDaR. Slides from the meeting will be posted on the RaDaR website.

3. **How can I learn about additional changes to the GRDR (now RaDaR)?**

As plans for the registries project continue to evolve, they will be posted on the registry project website and through other communication vehicles (such as NCATS notices), where applicable.

To receive NCATS and ORDR email notifications, please register with "Connect with NCATS" by clicking here and here.

You may also sign up to receive email updates directly on the new RaDaR website. The RaDaR website will be available at: https://rarediseases.info.nih.gov/radar

4. **Which elements of the GRDR will remain the same?**

The following GRDR services will be maintained and are accessible through the RaDaR website (Please visit the new RaDaR webpage for up-to-date links to these services):

- **Informed Consent template**: The Informed Consent Template will remain available on the registry website.
- **Global Unique Identifier (GUID)** A Global Unique Identifier (GUID) is a random computer-generated alphanumeric code that is assigned to each patient in a registry. The GUID is unique to each research participant within a registry, and permits de-identification of patients and their data to safeguard patient privacy. NCATS will continue to provide complimentary software that rare disease registry owners can download in order to generate the GUID. To use the tool, registry owners should submit a request per instructions on the GUID website.
- **Common Data Elements**: CDEs from the GRDR Registry Model will remain available on the National Library of Medicine (NLM) website.

5. **Will registry data be available through the RaDaR**

Data access from selected registries performed under the initial POC projects for the GRDR are not available through NCATS. Researchers interested in accessing these data for investigational purposes should contact the patient-advocacy registry owners for data access to specific registry data.

6. **Will additional registries be initiated through RaDaR?**

No additional registries will be funded through the RaDaR. Facilitation for registries will be provided through best-practices, advice and tools that will be made available through the RaDaR.

7. **Will data from additional existing registries undergo data mapping?**

No, no additional data mapping exercises for registries are planned.

8. **For the patient registries that underwent data mapping through the 2015 supplemental grant to HMS DBMI, how will the registry owners obtain access to their data?**

For the patient advocacy group registries that underwent data mapping, these registry datasets have undergone an extensive review, cleaning and mapping process by an expert team under Paul Avillach at HMS DBMI. The data have been mapped to more than 20 ontologies, and exist in Oracle and CSV formats, all of these being important standards in medical research. The benefit to these groups is that the data are now in a more shareable format because the mapping process has resulted in their greater interoperably. The benefit to the rare diseases community at large is that the HMS team has been able to provide some useful feedback on how rare diseases registries could best be set up going forward.

The patient advocacy groups whose registries underwent data mapping as part of this exercise should contact Paul Avillach at HMS DBMI to arrange for their mapped data to be sent to them. Please contact Dr. Avillach at: paul_avillach@hms.harvard.edu