Data Sharing FAQs

What is the Rare Disease Cures Accelerator-Data and Analytics Platform (RDCA-DAP®) and what does it do?

The Rare Disease Cures Accelerator-Data and Analytics Platform (RDCA-DAP) is an integrated database and analytics framework-designed to be used in building novel tools to accelerate drug development across rare diseases. It promotes the sharing of existing clinical data and encourages the standardization of the collection of new clinical data. By making such data available in a format suitable for analytics, RDCA-DAP accelerates the understanding of disease progression (including causes for variance in disease progression), clinical outcome measures and biomarkers, and facilitates the development of mathematical models of disease and innovative clinical trial designs. RDCA-DAP is positioned to generate solutions to drug development, which can be made publicly available to researchers in industry, government, regulatory agencies and academia. As such, the utility of the patient-level data will be maximized and used to develop tools that will be accessible to the community to optimize and accelerate drug development across rare diseases.

What data is in RDCA-DAP?

RDCA-DAP contains integrated data from multiple sources (clinical trials, natural history studies, registries, curated data from electronic health records etc.) across multiple rare diseases. Data is contributed from any source (initial sources include the IAMRARE ® databases from NORD and clinical trial and natural history data), and the custodians of these data have agreed to provide them to RDCA-DAP with defined provisions that are addressed in individual Data Use Agreements. Data ownership is maintained by contributor/custodian. As it becomes available, new datasets will be added to the database. All data is mapped to a common data standard and common terminology prior to integration. Datasets from a single source are not identifiable, however a user is able to identify the data type (clinical trial, registry, etc.).

Who has access to the data in RDCA-DAP?

RDCA-DAP allows tiered access to qualified researchers to the patient-level data dependent on the specific level of accessibility defined by the individual data custodians (people who share data with the platform). However, all the integrated data may be used to advance drug development tools. These tools may include models of disease progression, models to show the contribution of biomarkers to optimize clinical trial design, etc. These tools and results from analyses are made public.

How do I contribute data to RDCA-DAP?

Data may be contributed to RDCA-DAP from multiple sources. Custodians of the data need to sign a Data Use Agreement with C-Path and de-identify the data to ensure all consent and data privacy regulations are met and to designate how RDCA-DAP may use and share the data. Once this is completed, the data may be transferred from the contributor into RDCA-DAP. Files are uploaded into C-Path's encrypted file server. When the data are ready to be transferred, a

member of the C-Path data team will contact the custodian's data manager with instructions to complete this transfer. RDCA-DAP does not accept files transferred by any other means, including email. If the custodian has additional security or transfer requirements, C-Path will work with those guidelines provided the transfer process guarantees security. C-Path will collaborate with the data custodian to ensure that data is interpreted and mapped accurately.

Does the rare disease have to be approved by an RDCA-DAP committee before taking data for that disease? Will RDCA-DAP take data from a study in which it is the only data set in RDCA-DAP for that disease?
RDCA-DAP will accept data for any rare disease. The platform may actively seek data in specific disease areas identified as areas of need for analytics, but it will accept, integrate and make available any data that is shared with the platform. If only one dataset is available, and the custodian is willing to share it, then that data will be available as the only study for that disease. Our goal is to have multiple datasets for each disease. In the meantime, single datasets can be used for certain analyses and cross analyses with datasets in related diseases may be possible. If the custodian does not want their dataset to be identifiable, RDCA-DAP will not share the data externally until another dataset in the same disease area has been brought in.

Can an organization contribute data from an ongoing study, where prospective data is still being collected?
Data can be contributed throughout a study. For ongoing studies, RDCA-DAP will set up a system by which data can be transmitted at regular intervals (e.g., every two years) to keep the database up to date.

Who owns the data contributed to RDCA-DAP?
The original custodian of the data (the person or institution that shares the data with RDCA-DAP) retains ownership. They will dictate how RDCA-DAP may use the data through a legal agreement signed with the platform called a Data Use Agreement.

Is the data in RDCA-DAP de-identified? Can it be re-identified for future more in-depth analysis?
RDCA-DAP requires that all data is de-identified prior to sharing with the project to protect patient privacy. Our terms and conditions require that the user does not attempt to re-identify individuals in the database. In very small patient populations, however, it is difficult to fully de-identify the data, but by integrating worldwide data without standard identifiers (name, address, date of birth, etc.) we reduce the likelihood of identification. Any information in a dataset that could be identifiable will not be shared. If you need identifiable data, you will need to work with the person or group that collected the data originally.

What happens to data once it has been contributed to RDCA-DAP?
After data has been brought into C-Path, data managers will work with the data custodian and disease area experts to curate the data. Once the data is understood, it is tagged with common ontologies to allow searching across dataset. The data are made available as determined by the custodian's data sharing agreement and further curation and mapping to data standards to allow integration with other datasets may occur over time.

How secure is RDCA-DAP?
Data are encrypted during transfer and at rest, with access strictly controlled using a robust access management system and multi-factor authentication. Accounts with unique user ID's will be granted access to data sets in accordance with each set's governance structure and an audit trail logging of their activity. The platform itself is hosted in a high-security cloud infrastructure with continuous vulnerability scanning and full replication and backup. If data are removed from the platform, they will also be removed from all backup images of the data.

Will it adhere to HIPAA and GDPR and other regulations?
RDCA-DAP adheres to all applicable international data privacy and protection regulations and is fully GDPR-compliant. The following certifications, assessments and registrations are maintained by Aridhia Informatics in the Digital Research Environment (DRE) in which RDCA-DAP resides:

- ISO27001
- NHS Digital Security and Protection Toolkit Assessment
- Cyber Essentials Certification
- ICO Registration
- HITRUST Certification

Who can I contact with questions about the database?
An email can be sent to rdcadap@c-path.org with any questions. Contact information will also be provided on the RDCA-DAP website and within the platform itself.

Can contributors transfer patient-level data to C-Path, from studies in which the consent form prohibits the use of the study data for analyses other than the primary objectives of the study?
No. If the consent form prohibits the use of the study data for analyses other than the primary objectives of the study, the patient-level data cannot be transferred to C-Path.

Can contributors transfer patient-level data to C-Path from studies in which the consent form doesn't specifically prohibit the use of the study data for research other than the primary objectives of the study?
In some cases. Legal precedent in the U.S. has been set that, if the consent form does not specifically prohibit the use of the study data for research other than the primary objectives of the study, then the de-identified patient-level data can be used for research and, thus, can be transferred to C-Path.

Are there ways to share data with C-Path when the data must remain within the European Union (EU) or another governmental jurisdiction?
Because C-Path hosts data securely in a cloud environment, our servers can be physically located in many regions of the world. RDCA-DAP's data hub is based in the Netherlands. Where necessary and when the cloud service provider makes servers available, C-Path will store the data within a desired country or governmental jurisdiction but will still have the security, infrastructure and tools that are available in other regions. This means that C-Path can access shared data from the USA while hosting those data securely within in EU, for example, but may also aggregate those data with other data sources hosted elsewhere.

What steps does C-Path take to comply with GDPR?
C-Path pursues regulatory endorsement of drug development tools including biomarkers and disease progression models and submits supporting data to agencies such as the U.S. Food and Drug Administration (FDA) and the European Medicines Agency (EMA). These rigorous regulatory processes require full provenance of data along with assurance of security and compliance with privacy laws. Our de-identification practices help us comply with HIPAA and GDPR laws by removing PHI. Data contributors may also pseudonymize their patient-level data and destroy the key so that C-Path's copy of the data may be considered anonymized under GDPR.

For more FAQs specific to patients and patient organizations, visit
https://c-path.org/programs/rdca-dap/overview/resources/#faqs

FDA Acknowledgment