THERAPEUTIC ADVANCES in
Rare Disease

Review

Linking Angelman and dup15q data for expanded research (LADDER) database: a model for advancing research, clinical guidance, and therapeutic development for rare conditions

Sarah Nelson Potter, Elizabeth Reynolds, Katherine C. Okoniewski, Anne Edwards, Julia Gable, Christine Hill, Vesselina Bakalov, Stephanie Zentz, Carolyne Whiting, Emily Cheves, Katie Garbarini, Elizabeth Jalazo, Carrie Howell, Amanda Moore and Anne Wheeler

Ther Adv Rare Dis

2024, Vol. 5: 1–14

DOI: 10.1177/26330040241254122

***Abstract*:** Angelman syndrome (AS) and duplication 15q (dup15q) syndrome are rare neurogenetic conditions arising from a common locus on the long arm of chromosome 15. Individuals with both conditions share some clinical features (e.g. intellectual disability, epilepsy) and often require lifelong care. Disease-modifying therapies for both conditions are emerging, resulting in a significant need for a better understanding of the natural history of both AS and dup15q. Patient advocacy groups for both conditions recognized a need for a data repository that would link data on individuals from multiple sources to expand research, increase understanding of natural history, and accelerate the development of treatments, resulting in the Linking Angelman and Dup15q Data for Expanded Research (LADDER) Database. This paper describes the development and functionality of the LADDER Database – including challenges, lessons learned, and preliminary feasibility – and how it can be used as a model for other rare conditions.

Plain language summary

The LADDER database: a model for advancing research, clinical guidance, and therapeutic development for rare conditions

This paper describes the development and functionality of the Linking Angelman and Dup15q Data for Expanded Research (LADDER) Database, which is a data repository for two rare neurogenetic conditions: Angelman syndrome (AS) and duplication 15q (dup15q) syndrome. AS and dup15q syndrome arise from genetic abnormalities on chromosome 15 and share some clinical features (e.g. intellectual disability, epilepsy). LADDER was developed by patient advocacy organizations representing each condition in partnership with RTI International. LADDER links data on individuals from multiple sources to expand research, increase understanding of natural history, and accelerate the development of treatments for both AS and dup15q syndrome. The LADDER Database can be used as a model for expanding research and enhancing clinical trial readiness in other rare conditions.

Keywords: Angelman syndrome, data repository, dup15q syndrome, natural history, research, therapeutics

Received: 31 January 2024; revised manuscript accepted: 18 April 2024.

Correspondence to:
Sarah Nelson Potter
RTI International, 3040 E. Cornwallis Road, Research Triangle Park, NC 27709-2194, USA
snpotter@rti.org

Elizabeth Reynolds
Katherine C. Okoniewski
Anne Edwards
Julia Gable
Christine Hill
Vesselina Bakalov
Stephanie Zentz
Carolyne Whiting
Emily Cheves
Anne Wheeler
RTI International, Research Triangle Park, NC, USA

Katie Garbarini
Angelman Syndrome Foundation, Aurora, IL, USA

Dup15q Alliance, Matthews, NC, USA

Elizabeth Jalazo
Department of Pediatrics, The University of North Carolina School of Medicine, Chapel Hill, NC, USA

Carrie Howell
Dup15q Alliance, Matthews, NC, USA

Amanda Moore
Angelman Syndrome Foundation, Aurora, IL, USA

Patients with rare conditions face multiple unique challenges regarding diagnostic, prognostic, and therapeutic development processes. Most of these challenges originate from a lack of centralized knowledge about phenotypes and effective treatments, a natural result of the rarity of the conditions. Patient advocacy groups (PAGs) representing rare conditions recognize a profound need to connect patients with specialized medical centers, collect standardized clinician-reported outcomes, actively engage affected families in the research process, and provide a centralized platform for data to streamline research efforts. One example of a collaborative effort led by two rare disease PAGs – the Angelman Syndrome Foundation (ASF) and the Dup15q Alliance (Alliance) – to address these issues is the Linking Angelman and Dup15q Data for Expanded Research (LADDER) Learning Network (LLN). The LLN is a clinical and research network designed to help educate providers, connect families to specialized clinicians, integrate retrospective and prospective data to help improve knowledge, and build a foundation for natural history data and future recruitment to support clinical trials. Within the LLN is the LADDER Database, a centralized data repository developed to bring together multiple sources of information to streamline data and knowledge about these two rare disorders. The goal of the LADDER data repository is to facilitate a broad and deep understanding of each disorder and provide information for critical stakeholders working to develop and refine effective treatment protocols.

In the current paper, we describe the broad phenotypes and clinical needs of Angelman and duplication 15q (dup15q) syndromes, identify challenges facing rare disease communities, and highlight the roles of PAGs in developing research infrastructures to address barriers to treatment. The overall goal of the paper is to outline the development and functionality of the LADDER Database and explore how this data repository can be used as a foundation and working model for clinicians, researchers, advocates, and families in other rare conditions.

Angelman and dup15q syndromes

Angelman and dup15q syndromes are rare neurogenetic conditions, each of which affects approximately 1 in 15,000 individuals.[1–3] Both disorders arise from a common locus on the long arm of chromosome 15, at 15q11.2-13.1. Angelman syndrome (AS) is typically caused by a loss of maternally derived material at 15q11-q13, while dup15q syndrome is caused by a duplication of genetic material at 15q11-q13.[4] Additional less common genetic mechanisms in this region can also result in AS or dup15q syndrome.

AS and dup15q syndrome share some clinical features, including intellectual disability, speech and language delays, ataxia, sleep disturbances, and epilepsy.[5,6] Angelman-specific clinical features include absent or extremely limited verbal speech and happy affect.[4,7] Dup15q syndrome presents with a wider clinical spectrum of developmental challenges and is highly comorbid with autism spectrum disorder.[4]

Individuals with these conditions often have extremely limited independence and may require lifelong supervision and assistance with daily activities.[8,9] Therefore, the burden on caregivers of individuals with AS, dup15q, and associated conditions is high.[10] Caregivers of individuals with AS report sleep problems, anxiety, and physical pain.[9] In addition, families frequently experience financial problems resulting from the costs of therapies, specialized treatments, and medical devices, as well as compromised educational and/or employment opportunities due to the need to be full-time caregivers.[11,12]

There are currently no treatments for either condition that are approved by the Food and Drug Administration. Standards of care in the management of AS and dup15q include feeding support, epilepsy management, referrals to Early Intervention for those identified before 3 years of age, and comprehensive dynamic team monitoring across the lifespan.[13,14] These treatments are supportive and not curative. However, disease-modifying therapies for both conditions are emerging,[15,16] resulting in a need for more comprehensive natural history studies and streamlining of data in preparation for clinical trials.

Challenges facing rare disease communities

There are significant challenges associated with developing disease-modifying therapies for rare diseases like AS and dup15q. Challenges include fragmented research, limited collaboration

between interested parties, small and heterogeneous patient populations, and a lack of understanding of the disease's natural history.[17,18] Furthermore, bringing a new drug to market can take over a decade and cost over $2.5 billion.[19] Frequently in the rare disease sphere, research and data collection are siloed, stakeholders may be reluctant to share data and outcomes, and there are limited opportunities for stakeholders to effectively communicate and collaborate.[20] In addition, patients are geographically dispersed and may not live near specialized clinical research centers; even specialized physicians may only see a few patients per year.[18,21] In turn, biopharmaceutical companies have challenges recruiting participants, determining trial endpoints, and selecting clinical trial locations.

PAGs and research infrastructures

To overcome these challenges, PAGs have taken the lead in developing collaborative research networks, actively engaging patients in the research process, and improving coordination among stakeholders.[19,22] The creation, management, and maintenance of research infrastructure varies, but can include patient-reported outcome registries, data collection from clinical centers, genetic record curation, and electronic health record (EHR) integration. Gathering patient data is necessary to understand the natural history and trajectory of the disease, de-risk drug development, identify appropriate endpoints and valid outcome measures, and gain valuable caregiver perspective on decision-making.[20]

There are disadvantages and advantages to each type of research infrastructure approach (see Table 1). Altogether, these types of infrastructures can contribute to the larger goal of understanding disease trajectory, but alone may be insufficient to overcome the challenges faced by rare disease communities or to provide the breadth of information needed for clinical trials in their efforts to find treatments and cures. In the development of the LADDER Database, the advantages and disadvantages of previous efforts were considered along with key stakeholder perspectives and existing data sources to facilitate a comprehensive infrastructure that meets the varied needs of families, clinicians, researchers, and industry.

The LADDER database

Conceptualization and formation

PAG leaders, expert clinicians, and established researchers from the AS and dup15q communities recognized that established research infrastructures could contribute to the broad goals of understanding disease trajectory and advancing clinical trial preparedness, but alone were insufficient to overcome the challenges faced by their communities to find treatments and cures. Thus, stakeholders aimed to develop a comprehensive platform and system to link data on individuals with AS or dup15q syndrome with the ultimate aims of improving understanding of both disorders, maximizing research potential, and accelerating the development of interventions and treatments.

The LADDER Database originated with a growing awareness, particularly in the AS community, of the need for an expanded clinical network to serve patients and prepare the community for upcoming clinical trials. The process of expanding a formal clinic network highlighted the need for standardized data collection protocols that could be used to increase knowledge of the condition, identify priority clinical needs, and inform the development of clinical guidelines to ensure all patients receive the highest quality care regardless of where they seek healthcare. Furthermore, as preclinical findings began to point toward the possibility of disease-modifying therapeutics, the need to develop a functional research infrastructure became a priority.

To address these needs in the community, the ASF and Alliance developed a partnership to maximize resources and expand the clinic network. In 2019, the ASF and Alliance began working with researchers at RTI International to support the development of a database to synergize existing data sets in both communities and streamline future data collection from multiple sources. An initial meeting between the key stakeholders was held, including members of the funding PAGs, clinicians with expertise in one or both conditions, investigators from condition-specific natural history studies, and existing patient registries. This meeting resulted in the following goals for the LADDER Database:

Table 1. Existing PAG-developed data infrastructures: disadvantages and advantages.

Research infrastructure spearheaded by PAGs	Disadvantages	Advantages
PROs (i.e. patient or caregiver providing data *via* an online registry platform) Examples: Kuo *et al*.[23] Napier *et al*.[24] Reynolds *et al*.[25] Zilber *et al*.[26]	Not linked to clinician-reported data; challenging to validate Use of non-standardized measures Not always for 'research' (e.g. contact registry; no IRB oversight) Multiple registries for similar disorders The decision to share data for others to access is up to the registry; collected data can be siloed if there is an unwillingness to share Must coordinate with researchers to clean and harmonize data; foundations may not be familiar with data cleaning	Engages the patient community in research; patient communities are active participants Possibility of longitudinal data collection Inexpensive compared to clinical data collection Growing acceptance by the FDA to use PRO data No travel burden to families as forms can be accessed/completed virtually Families can access their data Can be spearheaded by PAGs (close relationship with the patient community; trust is important for recruitment) Flexibility and scaling for data collection Defines community for future clinical trial recruitment
Clinician-reported outcomes (e.g. natural history studies) Examples: Barca *et al*.[27] Pechmann *et al*.[28] Rummey *et al*.[29]	Data collection may not be standardized across sites; or only single-site data collection Expensive; frequently involves manual data entry by physicians and/or research coordinators Time intensive; can take years to develop study protocol, determine endpoints, recruit participants Frequently located in major cities; burden on families to coordinate appointments, travel Challenging to recruit families Data collection may be kept only at/for the site (i.e. siloed data) Possible incentives to *not* share data (financial, publications, etc.) Limited funding/infrastructure for analysis and dissemination of findings	Specialized care; physicians trained and knowledgeable about the rare disease Front-line observation of areas of strength and need Knowledge of clinical measures used to capture the profile of individuals; can support the selection of clinical trial endpoints Centers can become future sites for clinical trials (educated physicians, available families)
Genetic data curation Examples: Bladen *et al*.[30] Krishnaraj *et al*.[31] Strande *et al*.[32]	Not always linked to clinical data; hard to interpret phenotype/genotype relationship Challenging to recruit participants; need to ensure privacy Testing could be incomplete (e.g. no parent-origin testing)	Allows grouping by genetic subtypes
EHR Integration Examples: Brimble *et al*.[33] Brown *et al*.[34]	Consenting and integration challenges Need clear data elements to extract Need expertise in EHR data formats Privacy and security considerations for storing data The cost of the tool might be prohibitive for rare disease communities	Rare disease patients: small patient population but each individual has an extensive medical history Reduces the burden on families and physicians to report data Can include both retrospective and prospective data 21st Century Cures Act[35] standardized EHR data Potential ability to use EHR data as a historical control arm in clinical trials for rare diseases per FDA guidance

EHR, electronic health record; FDA, Food and Drug Administration; IRB, Institutional Review Board; PAG, patient advocacy group; PRO, patient-reported outcome.

1. Collect, store, and harmonize *existing* data sets on individuals with AS and dup15q. Bring together existing data sets from previous research studies, PAG registries, clinic visits, and clinical trials, and when possible, connect all data sources to a single participant. To be integrated, data need to be cleaned and transferred along with pertinent codebooks and database descriptions for integration into the larger data repository. To be linked, caregiver consent is required. Even with consent, some data could be shared with identifiers (e.g. date of birth) that could be used to help with linking, whereas other data sets would need to remain completely de-identified.
2. Collect, store, and harmonize *new* data sets on individuals with AS and dup15q. Develop new patient surveys, establish scheduled transfer of data from active research studies, request yearly update information for dup15q registry participants, and streamline data collection for network clinics. This goal is achieved in part through a core research study led by the LADDER Database which helps to support the standardization of data collected at clinic sites during clinic visits and natural history study visits across multiple locations.
3. Support and involve families in research. Enable families to be active research participants in secondary studies and provide user-friendly, accessible information so that they can learn from the data as it is collected.
4. Facilitate communication and collaboration among stakeholders. Bring PAGs, pharmaceutical representatives, clinicians, families, and researchers to a shared table to discuss research goals and active clinical trial progress. Encourage the sharing of information to work against siloed data collection and limited data sharing.

Overall, each of these goals supported the ultimate purpose of the data repository: to support researchers in their efforts to better understand these disorders and pursue novel therapies. The LADDER Database was formally established, with an equitable data-sharing policy that protects research participants and is simultaneously flexible to ensure maximum use of data and prevent data siloes.

Governance and funding

The collaborating PAGs contracted early in the process with investigators at RTI International, an independent nonprofit research institute dedicated to improving the human condition. RTI serves as the data coordinating center, providing the secure housing of data, oversight of regulatory and ethical aspects, data cleaning and harmonization, and management of adjacent patient- and clinic-facing research studies utilizing LADDER Database participants or data sets. Researchers with expertise in neurogenetic conditions, clinical trial preparation, data harmonization, data analysis, and outcome measure development comprise the core team overseeing the LADDER Database.

The ASF and Alliance, serving both as core funders and primary stakeholders, maintain ongoing representation and decision-making power for the LADDER Database. The mission of the ASF is to 'advance the awareness and treatment of Angelman syndrome through education and information, research, and direct support for individuals with AS, their families, and other parties'.[36] Similarly, the Dup15 Alliance seeks to 'empower individuals living with dup15q syndrome and other related rare diseases to reach their full potential by advancing breakthrough research and life-changing treatment, supporting families affected by dup15q, and promoting advocacy'.[37]

Since its conception, the LADDER Database team has worked to utilize the expertise of the governing parties in a collaborative and forward-moving framework. Meetings are held biweekly to provide updates on the progress of project goals, troubleshoot any problems, allow stakeholders to conceptualize new ventures for the project, and facilitate discussions about engagement and broader participation in the Database. Both the ASF and Dup15q Alliance fund the LADDER Database with their contributions made from pharmaceutical companies, private donors, and fundraising initiatives. The funding cycle runs on a fiscal year, with yearly budgets determined after mutual decision-making about project goals for the upcoming year.

Stakeholder engagement

The broader LLN aims to bring all critical parties (i.e. patients, caregivers, clinicians, and researchers) in the AS and dup15q communities under a comprehensive umbrella of education, clinical care provision, research opportunities, and advocacy. Integrating these parties into the LADDER Database has been a critical component of its success.

Researchers. Researchers who have collected data on individuals with AS or dup15q can contribute data to, or request data from, the LADDER Database. Data contributors maintain primary ownership of their data and are provided a seat on the data access committee (DAC), allowing them to maintain control over when and how their data are shared. More details regarding data access, sharing, and management are provided below.

Families. Patient and family-level participation is critical for the maintenance of the LADDER Database. Outreach and engagement with the AS and dup15q communities are facilitated through social media, attendance and participation at family-oriented conferences, solicitation through curated listservs, and establishment and support of LLN clinical sites. By participating in the LADDER Database, families have access to secondary studies and family-friendly resources, a streamlined registry response portal (for dup15q), and a centralized enrollment point for future clinical and research opportunities.

Clinical sites. The LADDER Database hosts a cornerstone project that works to facilitate engagement and participation by clinical network sites. The LADDER Database Clinical Needs Study (CNS) provides a streamlined intake form for parents/caregivers to complete upon establishing care at an ASF or Dup15q Clinic. The CNS caregiver form covers relevant medical and developmental domains, allowing the clinical team to focus on the most urgent clinical needs, reducing time spent in the clinic for both families and providers. Clinicians can orient and prepare for a clinical visit with both historical and current information on a patient through forms set up in an electronic database. In addition, the CNS provides clinician report forms, allowing for entry of patient information collected during visits (e.g. visit notes) as well as templates for commonly used assessments (e.g. standardized cognitive assessments, behavioral rating forms). Potential future integration with medical records at the sites is being considered.

Infrastructure

Development. The structure of the LADDER Database was developed with a stakeholder perspective on strengths and challenges identified through previous data collection/research infrastructure practices. Formative steps within the first year of the Database included (1) the development of a comprehensive, user-friendly website to provide information about the LADDER Database to researchers, clinicians, and families; (2) the establishment of an easily accessible account dashboard for parents/caregivers to enroll in the Database and access survey links for secondary studies; (3) the configuration of a searchable data dictionary and a data dashboard that provide information on data elements and those who have enrolled; and (4) discussions with clinical researchers and clinicians about standardizing forms of clinical utility as well as those from other large research studies [e.g. Angelman Syndrome Natural History Study[38]; Global Angelman Syndrome Registry (GASR)[39]].

Website and participant portal. The website is a centralized gateway that provides information about the LADDER Database for families, researchers, clinicians, and sponsors. The website is hosted in a scalable Amazon Web Services cloud environment that provides secure data access to database participants. To enroll in the LADDER Database, participants first need to create an account with the LADDER Database website. When a participant creates an account, they receive a confirmation email with a temporary password and a link to log in. After a participant logs in with their temporary password, they can complete the informed consent, Health Insurance Portability and Accountability Act (HIPAA)[40] Authorization, and enrollment forms. This process is in place to ensure the LADDER Database satisfies any HIPAA regulations and security requirements. The website is integrated with REDCap (Research Electronic Data Capture) to retrieve participant data from the REDCap server, and individual survey links are displayed on the participant dashboard after they log in. REDCap is a secure, web-based software platform designed

Figure 1. LADDER data dashboard.
LADDER, Linking Angelman and Dup15q Data for Expanded Research.

to support data capture for research studies.[41,42] REDCap provides (1) an intuitive interface for validated data capture; (2) audit trails for tracking data manipulation and export procedures; (3) automated export procedures for seamless data downloads to common statistical packages; and (4) procedures for data integration and interoperability with external sources.

The website is frequently updated to reflect changes in the survey requirements when new surveys are added, or existing surveys are closed. The website also supports the CNS by associating participants with a clinic *via* a clinic-specific link. Once a participant is linked with a clinic, they can consent and enroll in the LADDER Database (if they have not done so already) and complete the CNS forms specific to that clinic.

Data dashboard. An important component of the website is the data dashboard (see Figure 1). The dashboard is integrated with REDCap *via* Application Programming Interface to display the number of participants who have enrolled in the LADDER Database in real time. The dashboard allows users to filter historic data by different criteria such as age, gender, and molecular subtype. Future updates to the dashboard will include additional real-time statistics on Database participants (by sex, age, race/ethnicity, molecular subtype, and geographic location), and clinical outcomes of interest (e.g. age of seizure onset,

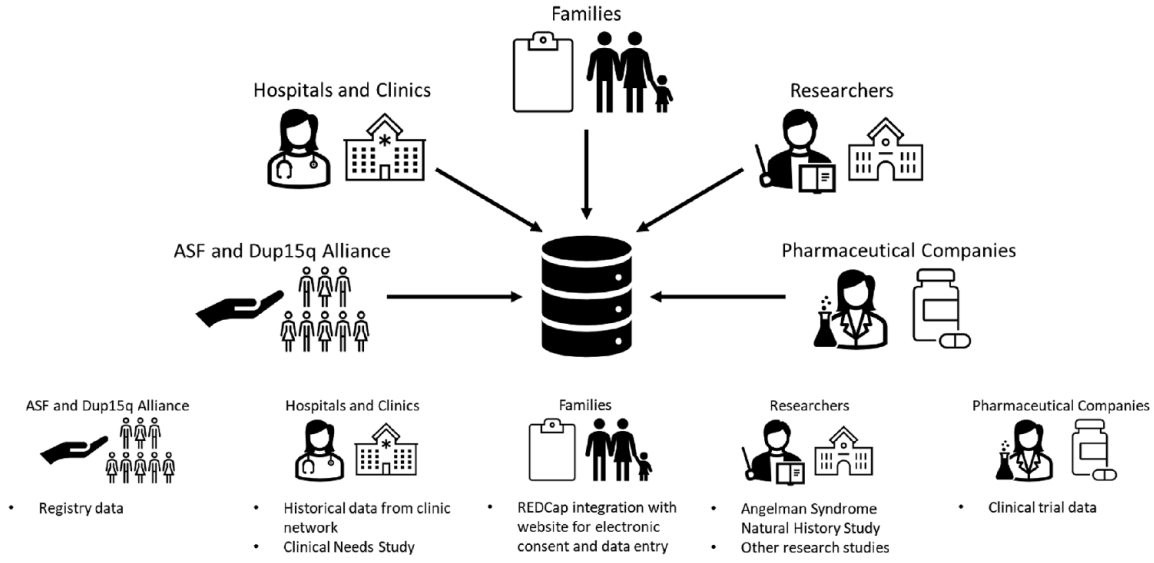

Figure 2. LADDER data sources.
LADDER, Linking Angelman and Dup15q Data for Expanded Research.

most frequently reported medications, medical problems).

Oversight and regulatory compliance

The Institutional Review Board (IRB) at RTI International provides regulatory consultation and oversight for the LADDER Database. Review of the original protocol was deemed non-human subjects research and reviews are conducted on an ad hoc basis as the Database evolves. Each secondary study offered to those enrolled in the Database goes under an individual review. RTI IRB provides oversight for most RTI studies. If the study is based on another institution, RTI IRB is asked to cede oversight to that institutional IRB. Approved secondary consent forms are implemented as instructed to maintain adequate human subjects' protections. RTI does not provide oversight for clinical data. As such, RTI engages with a private IRB for oversight of each clinical site that participates in the CNS.

Data access, management, and storage

Data are provided to LADDER by a variety of stakeholders and partners (see Figure 2). Researchers and industry partners who are known to have existing or emerging data sets on patients with AS or dup15q are invited to share their data with LADDER through direct requests, notices at national and international conferences, and *via* a link on the website. All data contributions are voluntary, and contributors can maintain control of how their data is used/shared through notifications when a data request is made and/or through participation on the DAC. Advantages for contributors to storing their data in the LADDER Database include having access to additional data for their research as well as options for long-term data storage, cleaning, and management.

The DAC, which is comprised of representatives from the funding PAGs, parent representatives from both conditions, and researchers who have contributed data, provides oversight of external data sharing. Formal proposals are submitted to RTI International through the LADDER website and the internal team disseminates the proposal to the DAC for review and consideration. The internal team reviews all requests and determines if a DAC review is needed based on the level of data access requested. There are three levels of data access provided by the LADDER Database. Level 1 allows any individual to preview LADDER Database content and data elements through a data dashboard (DAC review/approval not needed). Level 2 provides access to de-identified data sets for researchers. Level 3 supports the recruitment of participants for clinical trials or research studies and can involve survey development and distribution. All data access is free, and level 2 and level 3 requests require DAC review and approval. A monthly meeting is maintained

for the discussion of proposals. Alternatively, if upon review of the submitted research request DAC members feel as though an informal vote can be made virtually, one of the primary funding PAG leaders will begin the vote with tabulation recorded by internal team members at RTI.

The compliance team at RTI also provides oversight and review of how data travel into and out of the Database. A Data Use Agreement or Data Transfer Agreement (DUA and DTA, respectively) is executed after DAC approval is granted, ensuring that all security and regulatory guidelines are followed. Agreements include a description of the specific types of data requested, proposed use and analysis plans, data security policies, dissemination procedures, the time period for which data are requested, and a detailed outline of all terms and conditions.

Once the DUA/DTA is executed, contributed data sets are uploaded to a file transfer protocol (FTP) or sent by another secure transfer method. The data are usually in Excel or SAS files, or custom file formats (e.g. data from electroencephalogram [EEG] systems). After data are received, they are stored on the share drive on RTI servers. Contributed data sets with common elements are imported into SAS and integrated using SAS code to assign standard variable names and formats, align responses, create recodes, and convert items to standard units, as appropriate. Integrated data sets have a source variable to identify where each record or variable came from. Data sets that do not have common elements with other contributed data sets are stored in their original format. Quality checks are conducted by the data contributor before data are sent to the LADDER Database. The LADDER Database team then reviews data sets as they are received and addresses any questions with the contributor.

De-identified data are shared with requestors after approvals have been granted by the DAC and once a DUA/DTA is executed. Data sets are prepared according to the specifications in the DUA/DTA (e.g. a DUA might request data for a particular age group). Delivery includes SAS and Excel versions of the data, a list of data set contents, and a codebook containing frequencies and ranges. Data sets and documentation are combined into a password-protected zip file and uploaded to an FTP or transferred to the requester *via* another secure method. The requester receives an encrypted email with the name of the data file and instructions on how to access the FTP and download the data file. The password is sent to the requester in a separate encrypted email. The length of time sites maintain data is stated in the DUA, usually 1 year from receipt of the data. Sites can extend a DUA if they are interested in continuing to use the data beyond the term limit stated in the original DUA.

Dissemination and expansion

Secondary studies. The LADDER Database conducts secondary studies to gather information from enrolled participants on priorities identified by the community (e.g. epilepsy). These secondary studies are being completed by families in multiple countries and represent affected individuals across the lifespan. The data collected in these studies are shared with the community through social media, family-friendly reports, and peer-reviewed publications, and can also be used to inform treatment recommendations.

Social media. The LADDER Database launched its social media accounts, including Facebook and Instagram, in October 2022 to directly engage and reach the patient community. Content includes Frequently Asked Questions (e.g. Who can enroll in LADDER?), information about the types of data in LADDER (e.g. natural history studies, resulting data from clinical trials), and LADDER findings (e.g. publications and conference details). To share data and findings in a family-friendly manner, 'Saturday Stats' are shared (e.g. age of onset of seizures within each condition). In addition to disseminating information about LADDER, the social media accounts have been used to target recruitment of new participants (e.g. adult participants) and to encourage existing participants to complete secondary studies to complement their current profile (e.g. a survey to determine the prevalence of Lennox-Gastaut Syndrome among participants).

Family-friendly reports. A requirement for free access to data from LADDER is a brief family-friendly summary of any peer-reviewed manuscript that is published using LADDER data. These reports are available to view and download on the LADDER Database website and are also shared with the AS and dup15q communities through the ASF and Alliance. These reports help ensure that families have access to the

information that is generated from the data they share with LADDER.

Presentations and publications. To increase awareness of LADDER's existence and recruit both families as well as potential data contributors or users, members of the RTI team, with approval and support from the PAGs, have presented on LADDER at scientific symposiums and family conferences targeting AS or dup15q stakeholders. In addition, RTI provides analytic and manuscript preparation support for PAG-sponsored secondary studies as well as on a per-cost basis for external collaborators using LADDER data sets.

Challenges and lessons learned

We have encountered multiple challenges, many of which have been related to onboarding and conducting the CNS at multiple clinic sites across the United States. From a regulatory standpoint, because study oversight is maintained by a central IRB upon which each clinical site IRB relies, the IRB procedures are more complex than a single-site study and require multiple reviews and approvals from both the local and central IRBs. In addition, clinic staff completing the IRB submissions have varying amounts of experience with IRB with some requiring considerable guidance from LADDER staff. In addition, the onboarding process for clinic sites has oftentimes been prolonged due to understaffing, staff turnover, and institutional constraints, as well as the adoption of the CNS into existing clinic workflows.

Additional challenges include data cleaning logistics; some of the older contributed data sets have less documentation, making it more difficult to define and run data checks, identify common data elements from different studies, and define rules for combining variables. We have also encountered confusion about the many different resources, registries, and study opportunities within the AS and dup15q communities, which has at times limited buy-in from the communities (e.g. willingness to enroll or participate in secondary studies). We have had to adapt to these challenges and be flexible to continue to make progress toward our goals.

Preliminary feasibility

As of December 2023, over 460 caregivers of individuals with AS and 350 caregivers of individuals with dup15q have enrolled their child in LADDER. The following entities have contributed data to LADDER: ASF, Dup15q Alliance, Angelman Syndrome Natural History Study, Ovid Therapeutics, The University of North Carolina at Chapel Hill, Purdue University, and New York University. LADDER data elements include medical information and history (e.g. molecular diagnosis, physical exam, medications, medical problems, family history, behavioral history); measures of quality of life, stress, anxiety, behavior, sleep, and caregiver concerns; developmental assessments (e.g. Bayley Scales of Infant and Toddler Development); and other clinical assessments. In addition, 33 Level 2 and 3 Level 3 data access applications have been received, and 11 United States-based clinics have been onboarded for the CNS. Several strategies – including outreach *via* email campaigns and social media, presence at conferences, and the provision of resources (e.g. family-friendly reports) – are implemented to increase enrollment and participation in the CNS and secondary studies.

Future directions

Integration of EHRs. To be an effective data repository, LADDER strives to incorporate real-world data from multiple sources, including EHRs. There are multiple mechanisms to incorporate EHRs in data repositories. For example, researchers can access EHRs by requesting data and/or records from specific institutions or physicians. However, this review is dependent on physician and institutional approvals. Second, patients could scan or save documents they received from their appointments and share these PDFs. There are drawbacks to this approach, including incompleteness, time intensiveness for families, and significant time needed for researchers to review and collate results. Third, there are private businesses that offer EHR aggregation services for researchers, but limitations include siloed data, unclear data-sharing policies, and cost.

Within the last decade, there have been policy and technological shifts that have enabled a fourth mechanism to utilize EHRs in research: patient-mediated EHR integration. The 21st Century Cures Act mandated that patients be able to access, exchange, and use their electronic health information.[35,43] LADDER researchers are actively pursuing the use of a patient-facing tool

to collect and store EHR data, and preliminary survey and interview data show that LADDER participants are willing to share this type of data.

Expansion of international clinic network. As the LLN expands and international clinics join the Network to serve individuals with Angelman and dup15q syndromes across the globe, so does the interest in integration into the LADDER Database. As international clinics are set up, they are connected to Database team members. Then the standardized forms are shared along with codebooks to facilitate translation and integration into their electronic data systems. Once translated, forms are shared back with the Database to maintain a collection of translated materials and to ensure consistent and up-to-date use of forms.

Enhancing connectivity with GASR. The LADDER Database team, ASF, Foundation for Angelman Syndrome Therapeutics,[44] and GASR[39] are committed to data sharing to benefit the global AS community. There is an ongoing dialogue between these groups to facilitate data sharing, including the development of unique global identifiers for participants, the ability to request GASR data through LADDER, and the collection of standardized data elements from international clinic sites.

Conclusion

The LADDER Database was developed to link data on individuals with AS or dup15q to improve understanding of both disorders, increase research potential, and accelerate therapeutic development. Due to support from the funding PAGs, and collaborations with families, researchers, clinicians, and industry, the Database has demonstrated success in enrolling participants, spurring new research endeavors, launching secondary studies and the CNS at multiple clinic sites, and harmonizing data sets. The data repository is a substantial and important resource for the AS and dup15q communities and a model for other rare conditions.

Declarations

Ethics approval and consent to participate

RTI's Office of Research Protection determined that the Linking Angelman and Dup15q Data for Expanded Research (LADDER) Database (IRB STUDY00020896) was 'not research involving human subjects' and RTI IRB oversight was not required. Each secondary study offered to those enrolled in the Database goes under an individual review. As this is a review article describing the development and functionality of the LADDER Database consent to participate is not applicable.

Consent for publication

Not applicable.

Author contributions

Sarah Nelson Potter: Conceptualization; Methodology; Project administration; Writing – original draft; Writing – review & editing.

Elizabeth Reynolds: Conceptualization; Writing – original draft; Writing – review & editing.

Katherine C. Okoniewski: Conceptualization; Methodology; Project administration; Writing – original draft; Writing – review & editing.

Anne Edwards: Conceptualization; Writing – original draft; Writing – review & editing.

Julia Gable: Conceptualization; Methodology; Writing – original draft; Writing – review & editing.

Christine Hill: Conceptualization; Project administration; Writing – original draft; Writing – review & editing.

Vesselina Bakalov: Conceptualization; Writing – original draft; Writing – review & editing.

Stephanie Zentz: Conceptualization; Writing – original draft; Writing – review & editing.

Carolyne Whiting: Conceptualization; Writing – review & editing.

Emily Cheves: Conceptualization; Writing – review & editing

Katie Garbarini: Conceptualization; Writing – original draft; Writing – review & editing.

Elizabeth Jalazo: Conceptualization; Writing – original draft; Writing – review & editing.

Carrie Howell: Conceptualization; Writing – original draft; Writing – review & editing.

Amanda Moore: Conceptualization; Writing – original draft; Writing – review & editing.

Anne Wheeler: Conceptualization; Funding acquisition; Investigation; Methodology; Project administration; Supervision; Writing – original draft; Writing – review & editing.

Acknowledgements
We want to thank the families and researchers who have contributed to the LADDER Database.

Funding
The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the Angelman Syndrome Foundation (ASF) and Dup15q Alliance.

Competing interests
KG is a former employee of and owns stock in Invitae Laboratories. EJ is a paid consultant for both RTI International and the ASF. AW has received funding from the ASF, Foundation for Angelman Syndrome Therapeutics (FAST), Ovid Therapeutics, Ionis Pharmaceuticals, and F. Hoffmann-La Roche. The remaining authors declare that they have no relevant competing interests.

Availability of data and materials
Not applicable.

ORCID iD
Sarah Nelson Potter https://orcid.org/0000-0002-8640-8992

References

1. Angelman Syndrome Foundation. Facts about Angelman syndrome [Internet], https://www.angelman.org/wp-content/uploads/2019/12/facts_about_as_2009_3-19-10.pdf (2009, accessed December 13, 2023).
2. Isles AR, Ingason A, Lowther C, *et al.* Parental origin of interstitial duplications at 15q11.2-q13.3 in schizophrenia and neurodevelopmental disorders. *PLoS Genet* 2016; 12: e1005993.
3. Godler DE, Ling L, Gamage D, *et al.* Feasibility of screening for chromosome 15 imprinting disorders in 16 579 newborns by using a novel genomic workflow. *JAMA Netw Open* 2022; 5: e2141911.
4. Kalsner L and Chamberlain SJ. Prader-Willi, Angelman, and 15q11-q13 duplication syndromes. *Pediatr Clin North Am* 2015; 62: 587–606.
5. Bindels-de Heus K, Mous SE, Ten Hooven-Radstaake M, *et al.* An overview of health issues and development in a large clinical cohort of children with Angelman syndrome. *Am J Med Genet A* 2020; 182: 53–63.
6. DiStefano C, Gulsrud A, Huberty S, *et al.* Identification of a distinct developmental and behavioral profile in children with Dup15q syndrome. *J Neurodev Disord* 2016; 8: 19.
7. Bird LM. Angelman syndrome: review of clinical and molecular aspects. *Appl Clin Genet* 2014; 7: 93–104.
8. DiStefano C, Wilson RB, Hyde C, *et al.* Behavioral characterization of dup15q syndrome: toward meaningful endpoints for clinical trials. *Am J Med Genet A* 2020; 182: 71–84.
9. Wheeler AC, Sacco P and Cabo R. Unmet clinical needs and burden in Angelman syndrome: a review of the literature. *Orphanet J Rare Dis* 2017; 12: 164.
10. DiStefano C, Sadhwani A and Wheeler AC. Comprehensive assessment of individuals with significant levels of intellectual disability: challenges, strategies, and future directions. *Am J Intellect Dev Disabil* 2020; 125: 434–448.
11. Marsack-Topolewski CN and Church HL. Impact of caregiver burden on quality of life for parents of adult children with autism spectrum disorder. *Am J Intellect Dev Disabil* 2019; 124: 145–156.
12. Saunders BS, Tilford JM, Fussell JJ, *et al.* Financial and employment impact of intellectual disability on families of children with autism. *Fam Syst Health* 2015; 33: 36–45.
13. Duis J, Nespeca M, Summers J, *et al.* A multidisciplinary approach and consensus statement to establish standards of care for Angelman syndrome. *Mol Genet Genomic Med* 2022; 10: e1843.
14. Lusk L, Vogel-Farley V, DiStefano C, *et al.* Maternal 15q duplication syndrome. In: Adam MP, Feldman J, Mirzaa GM, *et al.* (eds.) *GeneReviews®*. University of Washington, Seattle, WA, 2016.
15. Hipp JF, Knoflach F, Comley R, *et al.* Basmisanil, a highly selective GABAA-α5 negative allosteric modulator: preclinical pharmacology and demonstration of functional target engagement in man. *Sci Rep* 2021; 11: 7700.

16. Markati T, Duis J and Servais L. Therapies in preclinical and clinical development for Angelman syndrome. *Expert Opin Investig Drugs* 2021; 30: 709–720.

17. Augustine EF, Adams HR and Mink JW. Clinical trials in rare disease: challenges and opportunities. *J Child Neurol* 2013; 28: 1142–1150.

18. Kempf L, Goldsmith JC and Temple R. Challenges of developing and conducting clinical trials in rare disorders. *Am J Med Genet A* 2018; 176: 773–783.

19. Kaufmann P, Pariser AR and Austin C. From scientific discovery to treatments for rare diseases – the view from the National Center for Advancing Translational Sciences – Office of Rare Diseases Research. *Orphanet J Rare Dis* 2018; 13: 196.

20. Denton N, Molloy M, Charleston S, *et al.* Data silos are undermining drug development and failing rare disease patients. *Orphanet J Rare Dis* 2021; 16: 161.

21. Dawkins HJS, Draghia-Akli R, Lasko P, *et al.* Progress in rare diseases research 2010-2016: an IRDiRC perspective. *Clin Transl Sci* 2018; 11: 11–20.

22. Rose SL. Patient advocacy organizations: institutional conflicts of interest, trust, and trustworthiness. *J Law Med Ethics* 2013; 41: 680–687.

23. Kuo A, Gomel R, Safer R, *et al.* Characteristics and outcomes reported by patients with primary sclerosing cholangitis through an online registry. *Clin Gastroenterol Hepatol* 2019; 17: 1372–1378.

24. Napier KR, Tones M, Simons C, *et al.* A web-based, patient driven registry for Angelman syndrome: the Global Angelman Syndrome Registry. *Orphanet J Rare Dis* 2017; 12: 134.

25. Reynolds E, Byrne M, Ganetzky R, *et al.* Pediatric single large-scale mtDNA deletion syndromes: the power of patient reported outcomes. *Mol Genet Metab* 2021; 134: 301–308.

26. Zilber S, Woleben K, Johnson SC, *et al.* Leigh syndrome global patient registry: uniting patients and researchers worldwide. *Orphanet J Rare Dis* 2023; 18: 264.

27. Barca E, Long Y, Cooley V, *et al.* Mitochondrial diseases in North America: an analysis of the NAMDC Registry. *Neurol Genet* 2020; 6: e402.

28. Pechmann A, Konig K, Bernert G, *et al.* SMArtCARE – a platform to collect real-life outcome data of patients with spinal muscular atrophy. *Orphanet J Rare Dis* 2019; 14: 18.

29. Rummey C, Corben LA, Delatycki M, *et al.* Natural history of Friedreich ataxia. *Neurology* 2022; 99: e1499–e510.

30. Bladen CL, Salgado D, Monges S, *et al.* The TREAT-NMD DMD Global database: analysis of more than 7,000 Duchenne muscular dystrophy mutations. *Hum Mutat* 2015; 36: 395–402.

31. Krishnaraj R, Ho G and Christodoulou J. RettBASE: Rett syndrome database update. *Hum Mutat* 2017; 38: 922–931.

32. Strande NT, Riggs ER, Buchanan AH, *et al.* Evaluating the clinical validity of gene-disease associations: an evidence-based framework developed by the clinical genome resource. *Am J Hum Genet* 2017; 100: 895–906.

33. Brimble E, Beek G, Wilson L, *et al.* PRO74 pilot study of a novel patient-facing medical records collection platform to support real world data generation in rare disease. *Value Health* 2021; 24: S211.

34. Brown TL, Nye KL and Porter BE. Growth and overall health of patients with SLC13A5 citrate transporter disorder. *Metabolites* 2021; 11: 746.

35. 21st Century Cures Act. H.R. 34, 114th Congress, https://www.gpo.gov/fdsys/pkg/BILLS-114hr34enr/pdf/BILLS-114hr34enr.pdf (2016, accessed December 13, 2023).

36. Angelman Syndrome Foundation-With you for the journey. [Internet], https://www.angelman.org (2023, accessed December 13, 2023).

37. Home - Dup15q Alliance. [Internet], https://www.dup15q.org/ (2021, accessed December 13, 2023).

38. ClinicalTrials.gov. 2000 Feb 29 – Identifier NCT04507997, Angelman Syndrome Natural History Study [Internet], Bethesda (MD): National Library of Medicine (US), https://clinicaltrials.gov/study/NCT04507997 (2020, accessed December 13, 2023).

39. Home - Global Angelman Syndrome Registry. [Internet], https://www.angelmanregistry.info/ (2023, accessed December 13, 2023).

40. GovInfo. Health Insurance Portability and Accountability Act. 104th Congress, https://www.govinfo.gov/app/details/PLAW-104publ191 (1996, accessed December 13, 2023).

41. Harris PA, Taylor R, Minor BL, *et al.* The REDCap consortium: building an international community of software platform partners. *J Biomed Inform* 2019; 95: 103208.

42. Harris PA, Taylor R, Thielke R, *et al.* Research electronic data capture (REDCap) – a metadata-driven methodology and workflow process for

providing translational research informatics support. *J Biomed Inform* 2009; 42: 377–381.

43. Lye CT, Forman HP, Daniel JG, *et al*. The 21st Century Cures Act and electronic health records one year later: will patients see the benefits? *J Am Med Inform Assoc* 2018; 25: 1218–1220.

44. Angelman Syndrome Research and Advocacy. Foundation for Angelman Syndrome Therapeutics. [Internet], https://www.cureangelman.org/ (2023, accessed December 13, 2023).

Appendix

Abbreviations

AS	Angelman syndrome
ASF	Angelman Syndrome Foundation
CNS	Clinical Needs Study
DAC	data access committee
DTA	Data Transfer Agreement
DUA	Data Use Agreement
dup15q	duplication 15q
EHR	electronic health record
FAST	Foundation for Angelman Syndrome Therapeutics
FDA	Food and Drug Administration
FTP	file transfer protocol
GASR	Global Angelman Syndrome Registry
HIPAA	Health Insurance Portability and Accountability Act
IRB	Institutional Review Board
LADDER	Linking Angelman and Dup15q Data for Expanded Research
LLN	LADDER Learning Network
PAGs	patient advocacy groups
PRO	patient-reported outcome
REDCap	Research Electronic Data Capture