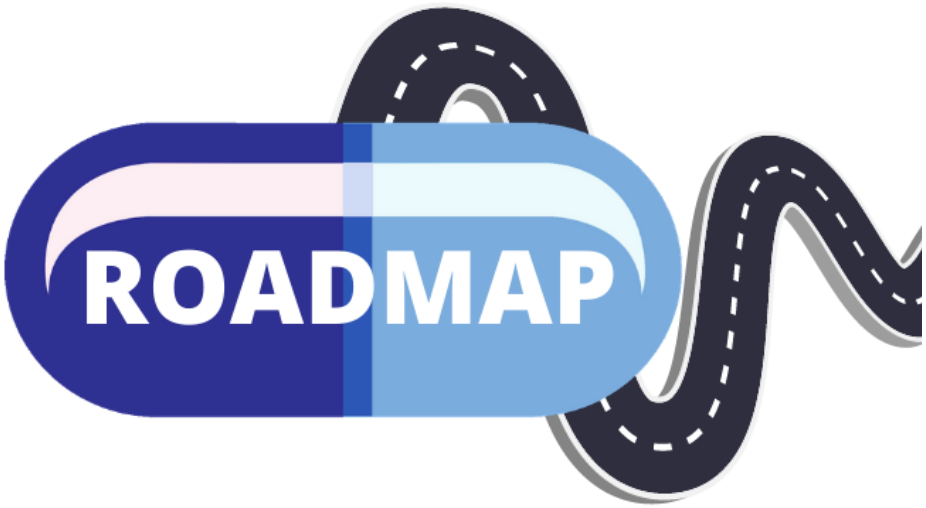
The ROADMAP Project



Conducted by the Castleman Disease Collaborative Network
Supported by the Chan Zuckerberg Initiative
In partnership with Every Cure

CDCN
CASTLEMAN DISEASE COLLABORATIVE NETWORK

CHAN ZUCKERBERG INITIATIVE

EVERY CURE

May 2021 - January 2023

Table of Contents

About the ROADMAP project

BACKGROUND

There are approximately 10,000 rare diseases affecting approximately 30,000,000 individuals in the U.S.A (1,2). Despite significant investments in resources over the last several decades, more than 90% of rare diseases do not have a single FDA-approved therapy (3). This fact highlights the failure of traditional new drug development to overcome the unique challenges faced by rare disease research and the need for alternative approaches.

Given that many diseases share the same underlying problems, the same treatments can often be used to treat multiple diseases. While this process is faster and less expensive than new drug development and repurposed treatments are sometimes found serendipitously, a number of barriers prevent this from happening systematically. Thus, repurposing existing FDA-approved drugs to treat rare diseases is a promising approach to help patients as quickly as possible. Similarly, drugs developed for one condition but proven ineffective or unsafe and thus never approved for the intended condition can also be used in another disease. Because the basic research and preliminary safety studies on the compound have already been completed, drug repurposing and repositioning is able to skip many steps and costs. Additionally, because the repurposed drug has already been proven safe and effective in another disease, it has a higher likelihood to succeed in clinical trials, compared to new, untested compounds (4).

For purposes of this project, the definition of drug repurposing is: "A process of research to identify potential treatments that are already FDA-approved or in development for one disease, for use in another disease by gathering data and analyzing efficacy in order to improve treatment guidelines and access". This includes the following:

- **Repurposing:** an FDA-approved drug used for another disease from the one that it is approved for.
- **Repositioning:** a drug with some safety and/or efficacy data in one disease and modifying its structure for use in another disease.
- **Reformulation:** a drug with some safety and/or efficacy data in one disease and modifying its method of administration or dose for use in another disease.
- **Rescue:** a drug that's not FDA-approved for any diseases due to complications in either safety and/or efficacy in the intended disease and used for another disease while maintaining the same structure and method of administration.

Due to the complexity of the steps involved for each type, the ROADMAP tool was created with a focus on repurposing drugs that are on the market, however, we refer to steps relevant to repositioning, reformulation or rescue throughout the tool.

There are a number of notable examples of drugs that have been repurposed to save rare disease patient lives, such as sildenafil for pulmonary arterial hypertension (originally for erectile dysfunction) (5), thalidomide for multiple myeloma (originally for leprosy) (6), and sirolimus for Castleman disease (originally for transplant rejection) (7).

Given the above, there is tremendous enthusiasm and interest in how to repurpose existing drugs for rare diseases, particularly from rare disease nonprofit organizations. However, the process and steps for how to effectively repurpose treatments is not clear. Through a partnership between the Castleman Disease Collaborative Network (CDCN), Chan Zuckerberg Initiative (CZI), and Every Cure, we launched the Repurposing Of All Drugs, Mapping All Paths (ROADMAP) project to identify the paths that can be taken to repurpose drugs, highlight the roles of various stakeholders, and centralize information on how to do this most effectively.

Acronyms that may be used throughout this document and the ROADMAP tool:
RDNP: rare disease nonprofit organization
CDCN: Castleman Disease Collaborative Network
SAB: Scientific Advisory Board
MAB: Medical Advisory Board
RFP: request for proposals
QOL: quality of life

METHODS

The Repurposing Of All Drugs, Mapping All Paths (ROADMAP) project aims to identify the paths that can be taken to repurpose drugs, highlight the roles of various stakeholders, and centralize information on how to do this most effectively through an interactive "ROADMAP" tool. This project is supported by a grant from the Chan Zuckerberg Initiative (CZI).

Figure 1: Six-phase project plan for ROADMAP project execution

In order to understand the most effective paths for rare disease drug repurposing, we needed to obtain data on the paths taken by rare disease stakeholders and the roadblocks they are facing. We began by identifying as many rare disease nonprofit organizations and potential stakeholders as possible, distributing a survey to these stakeholders, and performing in-depth interviews with a subgroup of representative stakeholders. Then, we

utilized this information along with our experiences from rare disease drug repurposing to build out the ROADMAP. We have separated out this approach into the following six steps:

1. Identifying and Characterizing stakeholders
2. Survey Execution
3. Survey Data Analysis
4. Interview Execution
5. Interview Data Analysis
6. Synthesis & Tool Development

This project was reviewed and approved by Advarra IRB under protocol number Pro00055201.

Citations

If you would like cite this project or the data available in the data explorer, please utilize this citation:

> Korsunska, A., Bolden, S. E., Repasky, M., Zuccato, M., Fajgenbaum, D. (2023) The ROADMAP Project. https://www.everycure.org/roadmap

NOTE: A paper describing the findings from this project will be submitted to a peer-review, open source journal soon, and will be made available for download under ROADMAP resources soon.

Identifying and Characterizing Stakeholders

Identifying all rare disease nonprofit organizations in the US

First, we set out to build a comprehensive list of all rare disease nonprofit organizations in the US. From here we hoped to gain an understanding of how many of them support research and drug repurposing in particular. Unfortunately, this list did not exist. We assembled a team of volunteers to combine several existing lists of organizations (i.e. NORD members, Global Genes members) and to perform additional searches to find more organizations. The final compiled list contained 982 organizations.

Characterizing rare disease nonprofit organizations

Next, a larger team of 70+ volunteers spent seven months (05/18/2021-12/16/2021) extracting data from the websites of these organizations, looking for:

- Whether the organization satisfied our inclusion criteria: US-based, registered 501(c)(3) nonprofit organization, focused on one or more rare diseases (one that affects less than 200,000 people in the US (2) and having an active website
- Basic organizational info: year of founding, name & contact information of founders

- Information on the organization's programs: conference, research agenda, biobank, registry, natural history study, etc.
- Whether the rare disease has treatment guidelines, or whether it is known that it is caused by a genetic mutation
- Any mention of drug repurposing, and if so - which drug(s) were pursued
- Any mention of partnerships or collaborations with other organizations

As a result of this exercise, **711** organizations were confirmed as rare disease nonprofit organizations in the US with active websites as of 12/2021. This dataset has been made available open-source here.

Gathering data on off label use from an established medical resource

We also onboarded a team of 10 extractors who had some medical expertise or medical educational background. We took a subset of 100 rare diseases from our ROADMAP survey data and conducted an extraction process of identifying how many rare diseases had off-label drugs as a part of the official treatment guidelines on a well-known and trusted electronic clinical resource tool for physicians. From **2/17/2022 to 7/14/2022**, the team extracted every drug mentioned on the rare disease treatment page regarding a specific rare disease in the list and coded the drug entry on its "context" and "status":

- **"Context"** refers to how the drug was mentioned on the rare disease treatment page. The options were "Context: Drug is recommended or listed as potentially helpful" and "Drug listed as does not work for this disease AND/OR no longer given AND/OR not recommended."
- **"Status"** refers to the drugs' approval status. The options were "FDA approved for another disease", "FDA approved for this rare disease" and "Not FDA approved for anything."

The goal of this data extraction was to understand the current state of rare disease treatment. A commonly used statistic is that 95% of rare diseases do not have an FDA-approved treatment. As far as we know, this statistic is never cited in any original research or dataset, so it is not clear how accurate it is. Furthermore, it does not tell us how many of these rare diseases have off-label (FDA approved for another disease, but not FDA approved for that rare disease) treatments available for them. This is crucial information for us to understand the state of rare disease treatment in the US and the need for drug repurposing.

Survey Execution

Survey Design

We designed a comprehensive survey using the Qualtrics platform, which included sections for several different stakeholder types: **(1)** rare disease nonprofit organization leaders, **(2)** rare disease patients, **(3)** rare disease patients' loved one *(parent, spouse, friend, sibling, etc. of a rare disease patient)*, **(4)** physicians who treat rare diseases, **(5)** rare disease researchers.

Any member of the leadership team of a US-based rare disease-focused nonprofit organization was able to participate in this research project. They were then invited to directly reach out to their US-based patient, loved

one, physician, and researcher network and invite them to take the survey. Since rare diseases predominantly affect children, we allowed participants under 18 to participate in the project through an adult loved one who provided informed consent and took the survey. Additionally, since many rare diseases cause physical and cognitive disabilities, we allowed those patients to also participate in the survey through a loved one, regardless of the patient's age. Also, we included an option for patients and loved ones to participate in Spanish, in order to be as inclusive as possible.

We focused our questions on potentially unique insights from each stakeholder:

- **Rare disease nonprofit organization leaders**: organizational characteristics (*age, funding, staff size, etc*.), level of resources (*biobank, patient registry, natural history study, scientific advisory board, etc.*), the rare disease state of research (*availability of treatment guidelines, diagnostic criteria, ICD code, biomarkers, etc.*), information on any FDA or off-label drugs being utilized, and drug repurposing experience, including challenges they've encountered, the support they need, and who was in their collaboration network.
- **Rare disease patients and their loved ones** (both stakeholders received the same set of questions): experience with off-label drugs for their rare disease and their general attitude towards and comfort level with drug repurposing.
- **Physicians**: the frequency of and factors that are taken into consideration when deciding to prescribe a drug off-label, off-label drug prescription practices for the rare diseases they treat, and levels of involvement in clinical trials for rare diseases.
- **Researchers**: involvement in research projects to advance drug repurposing for any rare diseases, what their motivations are, and what challenges they've encountered.

Preliminary Interviews

We conducted 10 preliminary semi-structured interviews with a selection of rare disease nonprofits, in order to receive feedback to refine the survey and project communication materials.

Survey Dissemination

We distributed the survey to all US-based rare disease organizations identified previously. Since this project is supported by a grant from the Chan Zuckerberg Initiative (CZI), the survey was also disseminated by the CZI leadership team to the organizations as a part of their Rare As One initiative.

The survey was officially launched on 9/29/2021 and stayed open for data collection until 1/6/2022**.** Because the survey included several sections and many of the respondents represented multiple stakeholders and were able to opt into multiple sections at once, we opted to include incomplete surveys at the end of the data collection period in order to gather the most data possible. In total, we received **1,923** total survey responses (completes and incompletes), which, after data cleaning and deduplication, included **723** total respondents (605 unique): **147** organization representatives, **340** patients, **170** loved ones, **23** physicians, and **43** researchers. Based on our search for rare disease nonprofit organizations, there are **711** active rare disease nonprofit organizations in the US. This means that we gathered data from approximately **20.7%** of the total sample , which is notable. Though there is an obvious selection bias to organizations who are interested or pursuing drug repurposing to take our

survey, review of the characteristics of these organizations suggests that we captured a variety of organizations.

Targeted Follow-up Survey

Of note, an error in the Qualtrics platform was identified in which a significant section of the patient/loved one survey was not shown to any participants. To address this, a follow up survey with these questions was sent to the 344 participants who consented to follow up. Among the **200** that responded, we were able to successfully match **191** to their original responses (See Merging Procedure below) and **156** met all inclusion criteria to be included in the full dataset for analysis.

Survey Data Analysis

The ROADMAP team did extensive data cleaning and analysis of the survey data, primarily using R studio (8), packages: tidyverse (9), stringr (10), dplyr (11), and googlesheets4 (12).

Inclusion Criteria Implementation & Data Cleaning

The data cleaning process consisted of five steps, which were executed concurrently:

1. Implementing inclusion criteria
2. Deduplication and merging
3. Name cleaning and translations
4. (Re-)classification and categorization
5. Data changes, corrections, and updates

1. Implementing inclusion criteria

Survey entries were removed from the final dataset for several reasons, including:

- → Organization was not a registered 501c3 nonprofit (**3**)
- → Organization does not support a rare disease (**2**)
- → Organization did not meet US location criteria (**155**)
- → Data submitted to Qualtrics after completion deadline (**6**)
- → Progress less than 10% survey completion (**399**)
- → Did not consent / did not answer consent question (**572**)
- → Did not answer any questions associated with stakeholder sections (**53**)
- → Data was a duplicate marked for removal (see "*Deduplication and merging"* for details) (**10**, not including RDNPs)

We applied each inclusion/exclusion criteria in the order presented here and updated the numbers of removals that were made after each step was applied to the dataset. This means that survey entries from later steps (e.g., "Didn't meet US location criteria") do **not** include entries from earlier steps (e.g., progress less than 10%).

In total, **1,324** entries (out of an original **1,929**) were removed due to not meeting inclusion criteria leaving **605** entries that qualified to remain in the final dataset.

2. Deduplication and merging

Deduplication Procedure

In some cases, the same individual completed the survey twice. To identify these cases, we looked for duplicate email addresses and/or IP addresses, as there was no other identifiable data collected via the survey. Then, we identified mismatches between their survey responses, if any, and manually reconciled these differences by combining information to make entries as complete as possible. We also resolved any contradictions and inconsistencies between entries that we identified.

There were some cases where multiple rare disease nonprofit representatives completed the survey for the same organization, causing this deduplication process to be more complex. To identify these cases, we looked for rare disease nonprofit name matches among those who opted into the associated section of the survey. Then, we compared the completion rates on these entries; entries with higher progress were automatically selected over entries with lower progress. In cases where the completion rate of each entry was equal (e.g., 100%), we compared the titles of each representative. In cases where one entry's representative had seniority (founder, CEO, executive director), we kept that entry. In a select handful of cases where both the progress and participant hierarchical standing were equal, we followed the same process described prior and manually reconciled any differences identified.

Table 1: Overview of deduplication process

Type of duplicate	Variables of interest	Identification method	Deduplication process
Cases where the same individual completed the survey twice	Question Q13 (email address) + IP addresses	Look for duplicate email and/or IP addresses.	Keep data that is the same between the multiple entries; reconcile differences between entries with contradicting information.
Cases where representatives from the same rare disease nonprofit completed the rare disease nonprofit stakeholder section of the ROADMAP survey	Question Q10 (stakeholder survey selection) + question Q301 (rare disease nonprofit name) + question Q17 (title) + Progress	Look for rare disease nonprofit name matches among those who opted into the rare disease nonprofit stakeholder section of the survey.	If one entry is more complete than another, keep the entry that is more complete. If completion rate is equal, but one of the entries was completed by a more senior representative founder, CEO, executive director); Keep data that is the same between the multiple entries; reconcile differences between entries with contradicting information.

Merging procedure

The merging procedure involved combining the data from the follow up survey with the original survey data as described in a prior section. We began with **200** follow-up survey entries, and we successfully matched **191** of these entries to their corresponding ROADMAP survey entries. To do so, we compared the IP and email addresses from the original and follow-up surveys, and we merged entries where matches were found. We then isolated the data on which drugs patients from each survey reported taking. We compared groupings of drugs, and combined entries where unique matching pairs were found. Ultimately, we discarded 9 follow-up survey entries that could not be matched to the original survey data via IP/email addresses or drug names.

It should be noted that while **191** successful matches were made, only **156** entries were ultimately combined with their original ROADMAP survey data. This is because **35** of the follow-up survey entries were from respondents whose entries were removed from the dataset during the inclusion filtering process outlined in the *"Filtering based on inclusion criteria"* and *"Deduplication and merging"* sections.

3. Name cleaning and translations

Name cleaning

We reviewed and cleaned all manually-inputted job titles, specialties, organization names, disease names, and drug names. While each process varied in its scale (whereas there were only a few unique specialties to clean, there were hundreds of drug names to clean) and scope (e.g., drug/disease name cleaning efforts were extended to additional projects), the actual cleaning process was nearly identical across each cleaning activity.

Each step in the cleaning process required isolating unique names, and then developing a "key-pair" dictionary with a column for the original name, and a column for its (optional) replacement. Members of the ROADMAP team, including volunteers, would then review the original names, identify inaccuracies and inconsistencies, and generate a final list of accurate replacements. In some cases, data was marked for removal (e.g., if a disease listed was not rare, or a drug listed as FDA-approved turned out to not be FDA-approved). Any use of brand-name drugs was replaced with their generic name, and the names of pharmaceutical salts were removed as well. Acronyms were avoided whenever possible. Each dictionary was imported into R and used to clean the corresponding survey data.

Table 2: Example key-pair dictionary

Original drug name	Replacement drug name
Tylenol	Acetaminophen
Tyelnol	Acetaminophen
Acetaminophen	Acetaminophen
Acetaminophen (500mg)	Acetaminophen
I can't remember	REMOVE

Translations

Two participants who met our inclusion criteria completed the Spanish-language version of the survey. To ensure that their data would be represented in our analysis, we translated their survey responses into English. In cases where translations of participant-entered text were required, two Spanish speakers (one native speaker, and one non-native speaker) reviewed each translation for accuracy.

4. (Re-)classification and categorization

For the purposes of analysis and visualization, we created new variables that were used to classify/categorize existing data. Examples include, but are not limited to the following:

- → Creating new categories for existing questions based on responses provided in the "other" sections
- → Classifying disease names
- → Classifying drug repurposing stages, status, and progress
- → Classifying job titles
- → Creating new variables (e.g., indicating whether drugs are FDA-approved based on human review; indicating whether an organization is focused on multiple diseases)

Additionally, some response categories were renamed, merged, and/or removed based on team expertise and consensus.

5. Data changes, corrections, and updates

We have made changes, corrections, and updates to our data on an ongoing, case-by-case basis. While we made minor updates to our patient, loved one, researcher, and physician data, most of our updates have been made to our rare disease nonprofit survey data. For example, in the rare case when a drug status change was identified among our list of 147 rare disease nonprofits since (e.g., a new drug being targeted for repurposing; a drug receiving FDA approval), we updated all survey question data corresponding to that drug. In a few instances where participant clarification was required to resolve inconsistencies, we reached out directly to rare disease nonprofits for their assistance using the email addresses that they provided. In some of these cases - such as when the participants directly let us know that the drug was listed in error - we would remove all data related to that drug from our dataset. Except for these major changes, we did not make updates based on any other status updates that were discovered throughout the research and interview process, and the dataset provides a snapshot of the state of drug repurposing as of survey date completion and some of the data may be or soon will be outdated.

Data Analysis

The survey data analysis process mostly focused on:

1. **1.** Descriptive statistics
2. **2.** Categorization
3. **3.** Visualization

1. Descriptive statistics

We performed descriptive statistics analysis to report basic quantitative statistical data about the ROADMAP survey. Examples include frequencies, measures of central tendency, and correlations. Generally speaking, we summarized raw survey data in our descriptive analysis. In a handful of cases, however, we developed new variables and/or created aggregations of variables for the purposes of analysis. All descriptive statistics have been calculated using R and R studio(8). The overarching goal of this descriptive analysis has been to characterize the rare disease nonprofits who participated in the ROADMAP survey (their typologies; outcomes and endpoints; characteristics; resources; etc.). Below, we review the general process that we followed to visualize, report, and share our findings from the ROADMAP survey. We also explore in more detail how we used the descriptive data that we collected to answer essential questions about the drug repurposing process and the progress of RDNPs in their drug repurposing journeys.

Most of the descriptive data that we generated were presented through bar charts, pie charts, and word clouds, among other visualizations. We created rough drafts of these visualizations using R. We generally focused on showcasing aggregated frequencies of survey options. In certain cases, multiple questions were combined and aggregated; in other cases, individual questions were broken down into multiple charts. Once we narrowed our selection of visualizations and created summaries of our findings, we built out a set of interactive web pages showcasing these final visualizations and the insights gathered from them. The ROADMAP Project's Survey Insights, available to explore here, are one of the ways that rare disease nonprofit organizations, as well as any other interested stakeholders, can explore our dataset and use it to inform further research or generate new insights.

Figure 2: Snapshot of the ROADMAP Project's Survey Insights

Drug Repurposing Stages

In order to analyze the progress of rare disease nonprofits in the drug repurposing process, we categorized the various steps involved in this process into stages. The stages that we developed were based on several questions from the survey and included:

- **Early stage**: Securing funding, trying to secure researcher interest to pursue, testing existing drugs in mouse or other animal models, patient data collection
- **Clinical trials**: Recruiting patients for clinical trials, executing clinical trials stages I, II, III
- **Late stage**: Analyzing clinical trial data, submitting for FDA approval, receiving FDA Breakthrough designation

Additionally, we worked to develop several success outcome categories, which would include more than just official FDA approval and provide alternative ways to conceptualize and measure drug repurposing success for rare diseases.

- **Endpoints:**
    - FDA approval
    - Off-label use, with some subjective measure of benefit (Drug to provide significant reduction in symptoms, improvement in quality of life, increase life expectancy/decrease in mortality, cure of disease, prevention of relapse)
    - Off-label use, without any subjective measure of benefit
    - Unsuccessful (project abandoned or FDA approval denied)

Once we had identified and classified these stages, we began to formulate specific questions about our data. While our general interest was to assess the relationship between organizational characteristics, drug identification methods, and stage(s) of progress, specific research questions included (but were not limited to):

- How many organizations had how many drugs in what stages currently?
- How each drug was identified and what stage is it in currently?
- How/Whether each organization's characteristics are related to the drug identification method used?
- How/Whether the drug identification method is related to the various types of “success” endpoints?

To answer these questions, we developed cross-tabulations of our data, which involved breaking the data into subgroups in order to look for patterns, trends, or other noteworthy observations. To be clear, this was a purely descriptive undertaking: we did not perform any kind of inferential or predictive statistics using our data. Instead, we focused on reporting raw frequencies, proportions, and percentages, and we used these numbers to help describe, characterize, and summarize our drug repurposing data, as well as answer the specific research questions outlined above.

One challenge that we had to address in developing cross-tabulations was that we were asking questions about two different units of analysis: in one case, drugs being repurposed (**94** total); in another case, the organizations involved in drug repurposing (**40** total). To address this, we made two separate sets of cross-tabulations, each taking a different unit of analysis. Both sets of cross-tabulations included the aforementioned list of drug repurposing stages (early, clinical trials, late, etc.). When taking repurposed drugs as our unit of analysis, we summarized this information on a per-drug basis. When taking organizations as our unit of analysis, we summarized this information on a per-organization basis. This required “flattening” the data of organizations repurposing multiple drugs into single yes/no responses. We, therefore, decided that when looking at this data on an organizational level, any “yes” responses from an organization’s survey data would overwrite any “no”

responses, as our goal here was to understand whether they had reached a given stage at *any* point; we relied on the drug repurposing cross-tabulations for more fine-grained information about how many specific drug repurposing projects reached each stage.

When taking organizations as our unit of analysis, we compared our stage data with our data on organizational characteristics/resources. When taking repurposed drugs as our unit of analysis, we compared our data on stages with our data on drug identification methods. In the table below, drug repurposing stages are represented in the rows, while drug identification methods are represented in the columns. Each stage / drug identification method is broken into a "yes" and "no" subcategory, where "yes" indicates that a survey-taker explicitly answered "yes" to the corresponding survey question(s), and vice versa for "no" selections. We filtered out cases where the question was left completely blank (instead of inferring, for example, that this indicates a "no"), meaning that our marginal totals (the sum of all cells in a given 2x2 contingency table) do not always neatly sum to 94/40.

Each set of 4 gray highlighted squares in the snapshot below represents a single 2x2 contingency table that we could study individually and/or compare against other 2x2 contingency tables.

	medical data (y)	medical data (n)	literature (y)	literature (n)	similar diseases (y)	similar diseases (n)	pre/trans (y)	pre/trans (n)	HTDS (y)	HTDS (n)	AI/ML (y)	AI/ML (n)	other (y)	other (n)
STAGE														
early (y)	6	19	2	22	2	23	14	11	6	19	1	24	6	19
early (n)	14	34	11	37	20	28	38	10	11	37	1	47	7	41
clinical (y)	6	21	6	21	9	18	24	3	6	21	1	26	4	23
clinical (n)	14	32	8	38	13	33	38	18	11	35	1	45	9	37
late (y)	6	8	3	11	6	8	9	5	5	9	0	14	3	11
late (n)	14	45	11	48	16	43	43	16	12	47	2	57	10	49
OUTCOME														
fda (y)	2	3	0	5	3	2	5	0	1	4	0	5	0	5
fda (n)	23	63	16	70	26	60	61	25	20	66	2	84	15	71
off-label (y)	10	9	4	15	9	10	14	5	6	13	0	19	0	19
off-label (n)	13	43	9	47	15	41	41	15	14	42	2	54	10	46
unsuccessful (y)	0	10	2	11	4	9	9	4	3	10	0	13	2	11
unsuccessful (n)	22	56	14	64	25	53	57	21	18	60	2	76	13	65
indeterminate (y)	11	46	10	47	16	41	41	16	12	45	2	55	13	44
indeterminate (n)	14	20	6	28	13	21	25	9	9	25	0	34	2	32

	medical data (y)	medical data (n)	literature (y)	literature (n)	similar diseases (y)	similar diseases (n)	pre/trans (y)	pre/trans (n)	HTDS (y)	HTDS (n)	AI/ML (y)	AI/ML (n)	other (y)	other (n)
STAGE														
early (y)	8.22%	26.03%	2.78%	30.56%	2.74%	31.51%	19.18%	15.07%	8.22%	26.03%	1.37%	32.88%	8.22%	26.03%
early (n)	19.18%	46.58%	15.28%	51.39%	27.40%	38.36%	52.05%	13.70%	15.07%	50.68%	1.37%	64.38%	9.59%	56.16%
clinical (y)	8.22%	28.77%	8.22%	28.77%	12.33%	24.66%	28.92%	3.61%	8.22%	28.77%	1.37%	35.62%	5.48%	31.51%
clinical (n)	19.18%	43.84%	10.96%	52.05%	17.81%	45.21%	45.78%	21.69%	15.07%	47.95%	1.37%	61.64%	12.33%	50.68%
late (y)	8.22%	10.96%	4.11%	15.07%	8.22%	10.96%	12.33%	6.85%	6.85%	12.33%	0.00%	19.18%	4.11%	15.07%
late (n)	19.18%	61.64%	15.07%	65.75%	21.92%	58.90%	58.90%	21.92%	16.44%	64.38%	2.74%	78.08%	13.70%	67.12%
OUTCOME														
fda (y)	2.20%	3.30%	0.00%	5.49%	3.30%	2.20%	5.49%	0.00%	1.10%	4.40%	0.00%	5.49%	0.00%	5.49%
fda (n)	25.27%	69.23%	17.58%	76.92%	28.57%	65.93%	67.03%	27.47%	21.98%	72.53%	2.20%	92.31%	16.48%	78.02%
off-label (y)	13.33%	12.00%	5.33%	20.00%	12.00%	13.33%	18.67%	6.67%	8.00%	17.33%	0.00%	25.33%	0.00%	25.33%
off-label (n)	17.33%	57.33%	12.00%	62.67%	20.00%	54.67%	54.67%	20.00%	18.67%	56.00%	2.67%	72.00%	13.33%	61.33%
unsuccessful (y)	0.00%	11.36%	2.20%	12.09%	4.40%	9.89%	9.89%	4.40%	3.30%	10.99%	0.00%	14.29%	2.20%	12.09%
unsuccessful (n)	25.00%	63.64%	15.38%	70.33%	27.47%	58.24%	62.64%	23.08%	19.78%	65.93%	2.20%	83.52%	14.29%	71.43%
indeterminate (y)	12.09%	50.55%	10.99%	51.65%	17.58%	45.05%	45.05%	17.58%	13.19%	49.45%	2.20%	60.44%	14.29%	48.35%
indeterminate (n)	15.38%	21.98%	6.59%	30.77%	14.29%	23.08%	27.47%	9.89%	9.89%	27.47%	0.00%	37.36%	2.20%	35.16%

Figure 3: Snapshot of the contingency table for drug identification methods and where the drug is currently (top is raw values, bottom is proportions with highlighting to signify percentage ranges)

For a concrete example of how the above contingency table is read, we can "zoom in" on the top-left 2x2 contingency table, which displays information about two subgroups of drugs being repurposed: those in the "early" stage of drug repurposing and those that relied on "medical data" in the drug identification process.

Looking at the below table tells us the following information about the relationship between this specific stage and drug identification method:

- **6** repurposed drugs are currently in the early stage of drug repurposing and were identified using medical data
- **19** repurposed drugs are currently in the early stage of repurposing but were *not* identified using medical data
- **14** repurposed drugs were identified using medical data, but are *not* in the early stage of repurposing
- **34** repurposed drugs were *not* identified using medical data, and are *not* in the early stage of repurposing

STAGE	medical data (y)	medical data (n)
early (y)	6	19
early (n)	14	34

Figure 4: Example of a 2x2 contingency table (comparing drug repurposing stages with drug identification data)

2. Categorization

One of the important outcomes of the ROADMAP project was a comprehensive understanding of all the “menu” items in the drug repurposing process, namely all the steps involved and all the options that a researcher can pursue or a representative of a rare disease nonprofit can support. Based on our experience and expertise, the ROADMAP team implemented a preliminary list of options in the survey, but via user feedback from various stakeholders, we were able to build out a more detailed breakdown of both the stages involved in the drug repurposing process as well as the options that are available for selection within each step. This categorization was done through an ongoing manual, collaborative visualization exercise using the tool MIRO, which allows for mind mapping, timeline building and other visualization techniques in an infinite, collaborative workspace. We also utilized this approach to merge the interview insights with the existing survey data, which we will describe in a later section.

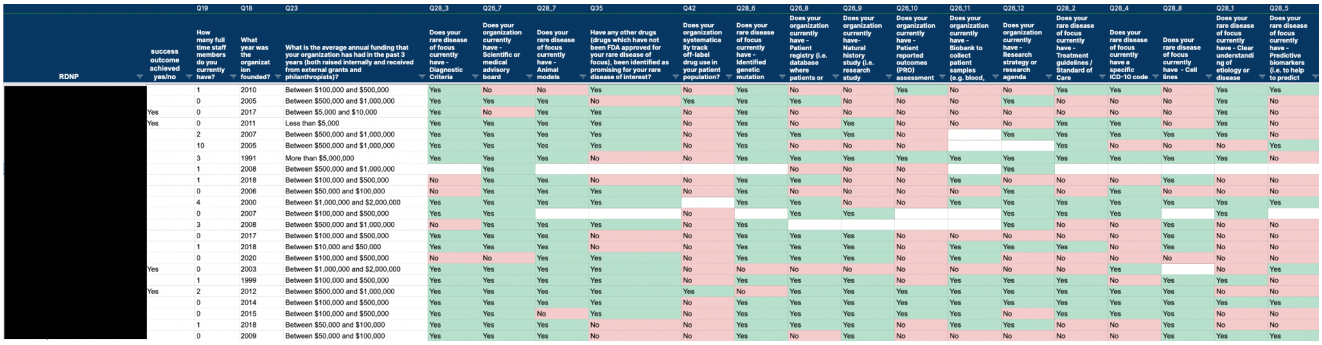
Figure 5: Illustration of the categorization processes done in MIRO

3. Network Visualization

One of the sections of the rare disease nonprofit organization representative survey focused exclusively on their collaboration network. In this section, the rare disease nonprofit organization representatives were asked to list their top collaborator organizations by name, and then specify what kind of collaboration activities they were engaged in with them. From this data, we also built a network dataset of nodes (organizations) and edges (the relationships between organizations). We visualized this network using a tool called Gephi (13). This allowed us to gain a better understanding of the types of relationships going on between different rare disease nonprofit organizations, as well as their external links to research institutions, pharmaceutical companies and government organizations.

Interview Execution

From the 147 rare disease nonprofit organizations in our survey data, we developed a typology of organizations based on various internal characteristics and drug repurposing experience.



RDNP	success outcome achieved yes/no	Q19 How many full time staff members do you currently have?	Q18 What year was the organization founded?	Q23 What is the average annual funding that your organization has had in the past 3 years (both raised internally and received from external grants and philanthropists)?	Q28_3 Does your rare disease of focus currently have - Diagnostic Criteria	Q26_7 Does your organization currently have - Scientific or medical advisory board	Q28_7 Does your rare disease of focus currently have - Animal models	Q35 Have any other drugs (drugs which have not been FDA approved for your rare disease of focus), been identified as promising for your rare disease of interest?	Q42 Does your organization systematically track off-label drug use in your patient population?	Q28_6 Does your rare disease of focus currently have - Identified genetic mutation	Q26_8 Does your organization currently have - Patient registry (i.e. database where patients or	Q28_9 Does your organization currently have- Natural history study (i.e. research study	Q26_10 Does your organization currently have - Patient reported outcomes (PRO) assessment	Q26_11 Does your organization currently have - Biobank to collect patient samples (e.g. blood,	Q26_12 Does your organization currently have - Research strategy or research agenda	Q28_2 Does your rare disease of focus currently have - Treatment guidelines / Standard of Care	Q28_4 Does your rare disease of focus currently have a specific ICD-10 code	Q28_8 Does your rare disease of focus currently have - Cell lines	Q28_1 Does your rare disease of focus currently have - Clear understanding of etiology or disease	Q28_5 Does your rare disease of focus currently have - Predictive biomarkers (i.e. to help to predict
[redacted]		1	2010	Between $100,000 and $500,000	Yes	No	No	Yes	No	Yes	No	No	Yes	No	No	Yes	Yes	No	Yes	Yes
		0	2005	Between $500,000 and $1,000,000	Yes	Yes	Yes	No	Yes	Yes	Yes	No	No	No	No	Yes	No	No	No	Yes	No
	Yes	0	2017	Between $5,000 and $10,000	Yes	No	Yes	Yes	No	Yes	No	No	No	No	No	No	No	No	Yes	No
	Yes	0	2011	Less than $5,000	Yes	Yes	Yes	Yes	No	Yes	No	Yes	No	No	No	Yes	Yes	No	Yes	No
		2	2007	Between $500,000 and $1,000,000	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	No		Yes	Yes	Yes	Yes	Yes	No
		10	2005	Between $500,000 and $1,000,000	Yes	Yes	Yes	Yes	No	Yes	No	No	No			Yes	No	No	No	Yes
		3	1991	More than $5,000,000	Yes	Yes	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No
		1	2008	Between $500,000 and $1,000,000		Yes					No	No	No		Yes					
		1	2018	Between $100,000 and $500,000	No	Yes	Yes	No	No	Yes	No	No	No	Yes	No	No	No	No	Yes	No	No
		0	2006	Between $50,000 and $100,000	No	Yes	Yes	Yes	No	Yes	No	No	No	No	Yes	No	Yes	No	No	No
		4	2000	Between $1,000,000 and $2,000,000	Yes	Yes	Yes	Yes		Yes	Yes	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
		0	2007	Between $100,000 and $500,000	Yes	Yes			No		Yes	Yes			Yes	Yes	Yes		Yes	
		3	2008	Between $500,000 and $1,000,000	No	Yes	Yes	Yes	No	Yes					Yes	No	No	Yes	No	No
		0	2017	Between $100,000 and $500,000	Yes	Yes	Yes	No	No	Yes	Yes	Yes	No	No	No	No	No	Yes	No	No
		1	2018	Between $10,000 and $50,000	Yes	Yes	Yes	No	No	Yes	Yes	Yes	No	Yes	Yes	Yes	No	Yes	No	No
		0	2020	Between $100,000 and $500,000	No	No	Yes	Yes	No	Yes	Yes	Yes	No	Yes	No	No	No	No	No	No
	Yes	0	2003	Between $1,000,000 and $2,000,000	Yes	Yes	Yes	Yes	No	No	No	No	No	No	No	No	No	Yes		No
		1	1999	Between $100,000 and $500,000	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	No	Yes	Yes	Yes	No	Yes	Yes	No
	Yes	2	2012	Between $500,000 and $1,000,000	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No
		0	2014	Between $100,000 and $500,000	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
		0	2015	Between $100,000 and $500,000	Yes	Yes	No	Yes	No	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	No	No
		1	2018	Between $50,000 and $100,000	Yes	Yes	Yes	No	No	Yes	Yes	Yes	No	Yes	Yes	No	No	Yes	No	No
		0	2009	Between $50,000 and $100,000	Yes	Yes	Yes	No	No	Yes	No	Yes	No	No	No	No	No	Yes	Yes	Yes

Figure 6: Snapshot of the typology of organizations for interviewee selection
[RDNP names hidden for data confidentiality]

Based on the typology, we selected **32** rare disease nonprofit organizations for participation in the interview phase of the project. We selected all **15** rare disease nonprofit organizations which selected at least one subjective success endpoint to have been met, and **17** organizations where no subjective success endpoints have been met yet. Out of these 17, we selected a sample from each stage of the repurposing process, and organizations that were comparable by other characteristics (age, funding level, staff size, etc.) We did comparisons of our selected organizations both internally (15 vs 17 organizations) as well as comparing them to the total 147, as well as those we selected for interviews vs those not selected (32 vs 115). Overall, we were satisfied with our selection, that it captured a good variability across all important characteristics.

		15 "met a success endpoint"	17 in "process/unsuccessful"	147 all
Funding	Less than $5,000	6.67%	0.00%	4.17%
	Between $5,000 and $10,000	6.67%	0.00%	4.86%
	Between $10,000 and $50,000	6.67%	6.25%	17.36%
	Between $50,000 and $100,000	0.00%	6.25%	10.42%
	Between $100,000 and $500,000	33.33%	50.00%	32.64%
	Between $500,000 and $1,000,000	13.33%	12.50%	13.19%
	Between $1,000,000 and $2,000,000	20.00%	12.50%	8.33%
	Between $2,000,000 and $5,000,000	6.67%	6.25%	4.17%
	More than $5,000,000	6.67%	6.25%	4.17%
Age	pre 2000	46.67%	12.50%	29.17%
	post 2000	53.33%	87.50%	76.87%
Size of staff	0 staff	40.00%	18.75%	30.82%
	1 staff	0.00%	56.25%	26.71%
	2-5 staff	26.67%	12.50%	21.92%
	5-10 staff	26.67%	12.50%	10.27%
	>10 staff	6.67%	0.00%	4.11%
Yes for Diagnostic Criteria		100.00%	93.75%	82.86%
Yes for SAB/MAB		93.33%	100.00%	92.31%
Yes for Animal models		93.33%	87.50%	78.20%
Yes for Off label drugs identified -		86.67%	76.92%	66.39%
No - do not systematically track off label use		60.00%	87.50%	87.77%
Yes for identified genetic mutation		73.33%	93.75%	82.35%
Yes for Patient registry		53.33%	75.00%	59.86%
Yes for natural history study		53.33%	75.00%	42.14%
Yes for Patient reported outcomes		42.86%	43.75%	24.82%
Yes for biobank		40.00%	56.25%	34.09%
Yes for research agenda		66.67%	87.50%	55.07%
Yes for treatment guidelines		57.14%	62.50%	51.45%
Yes for ICD code		71.43%	43.75%	47.45%
Yes for Cell lines		71.43%	93.75%	63.57%
Yes for Clear understanding of etiology or disease pathogenesis		46.67%	68.75%	49.28%
Yes for Predictive biomarkers		28.57%	25.00%	25.56%
Yes for FDA approved drugs		33.33%	12.5	29.50%

	Category	n = 147 # all	n = 25 # interviewed	n = 122 # not interviewed	n = 147 % all	n = 25 % interviewed	n = 122 % not interviewed
funding	Less than $5,000	6	1	5	4.08	4	4.1
	Between $5,000 and $10,000	7	1	6	4.76	4	4.92
	Between $10,000 and $50,000	25	2	23	17.01	8	18.85
	Between $50,000 and $100,000	15	0	15	10.2	0	12.3
	Between $100,000 and $500,000	48	10	38	32.65	40	31.15
	Between $500,000 and $1,000,000	19	2	17	12.93	8	13.93
	Between $1,000,000 and $2,000,000	12	6	6	8.16	24	4.92
	Between $2,000,000 and $5,000,000	6	2	4	4.08	8	3.28
	More than $5,000,000	6	1	5	4.08	4	4.1
age	2000 and prior	36	6	30	24.49	24	24.59
	post 2000	111	19	92	75.51	76	75.41
staff size	zero	52	6	46	35.37	24	37.7
	1 to 1.5	40	7	33	27.21	28	27.05
	2 to 5	32	6	26	21.77	24	21.31
	6 to 10	15	5	10	10.2	20	8.2
	11 or more	7	1	6	4.76	4	4.92
etc	Yes for Diagnostic Criteria	123	24	99	83.67	96	81.15
	Yes for SAB/MAB	136	24	112	92.52	96	91.8
	Yes for Animal models	118	23	95	80.27	92	77.87
	Yes for Off label drugs identified -	107	22	85	72.9	88	69.67
	No - do not systematically track off label	130	18	112	88.44	72	91.8
	Yes for identified genetic mutation	123	20	103	83.67	80	84.43
	Yes for Patient registry	90	14	76	61.22	56	62.3
	Yes for natural history study	66	15	51	44.9	60	41.8
	Yes for Patient reported outcomes	44	8	36	29.93	32	29.51
	Yes for biobank	60	10	50	40.82	40	40.98
	Yes for research agenda	85	19	66	57.82	76	54.1
	Yes for treatment guidelines	80	15	65	54.42	60	53.28
	Yes for ICD code	75	15	60	51.02	60	49.18
	Yes for Cell lines	100	21	79	68.03	84	64.75
	Yes for Clear understanding of etiology o	77	14	63	52.38	56	51.64
	Yes for Predictive biomarkers	48	6	42	32.65	24	34.43
	Yes for FDA approved drugs	49	6	43	33.33	24	35.25

Figures 7: Analysis of organizational subgroups

We then invited a member of the leadership team of all **32** organizations to participate via email, and **25** agreed to participate and were interviewed. The ROADMAP project lead conducted all 25 60-minute semi-structured interviews with them over the course of two months via ZOOM. The interview protocol focused on questions regarding their experience of drug repurposing, based on a preliminary framework consisting of four stages: **1.** Preparation, **2**. Drug Identification, **3.** Validations, **4.** Clinical Use and **5.** Optimal Utilization. Within each step, we generally discussed how they executed each step, how long each step took, how the decision was made as to what approach to take, what challenges they faced, what advice they have for other organizations, who they collaborated with, and what resources they utilized. In cases when organizations were in the process of

repurposing multiple drugs, the interviewees were asked to discuss the ones that were either the farthest along in the process or with which they had the most varied experiences so that on aggregate, the ROADMAP would contain insights from as many varied experiences as possible. Because of this, even though **25** organizations were interviewed, we were able to capture the experiences of repurposing **75** drugs.

Interview Data Analysis

The interview audio was transcribed and cleaned utilizing Otter.ai, with a final manual review from research assistants and project lead to guarantee the accuracy of transcript data. Then, the project team utilized a tool called Dedoose to conduct an iterative thematic analysis of the interview transcripts (14). The project lead identified and coded statements from the interview transcript that answered a research question and generated a list of themes from these codes; this resulted in 120 themes. The project lead and research assistant then synthesized the themes into five categories of findings: **1. General Description, 2. Insights, 3. Roadblocks, 4. Support Options and 5. Resources.** We then utilized a mind mapping software MIRO, as mentioned in previous sections, to identify and categorize data visually and collaboratively. This process was crucial in making sure our interview insights helped fill in any missing categories from our survey data, and provide additional advice or nuance that we were not able to capture via survey.

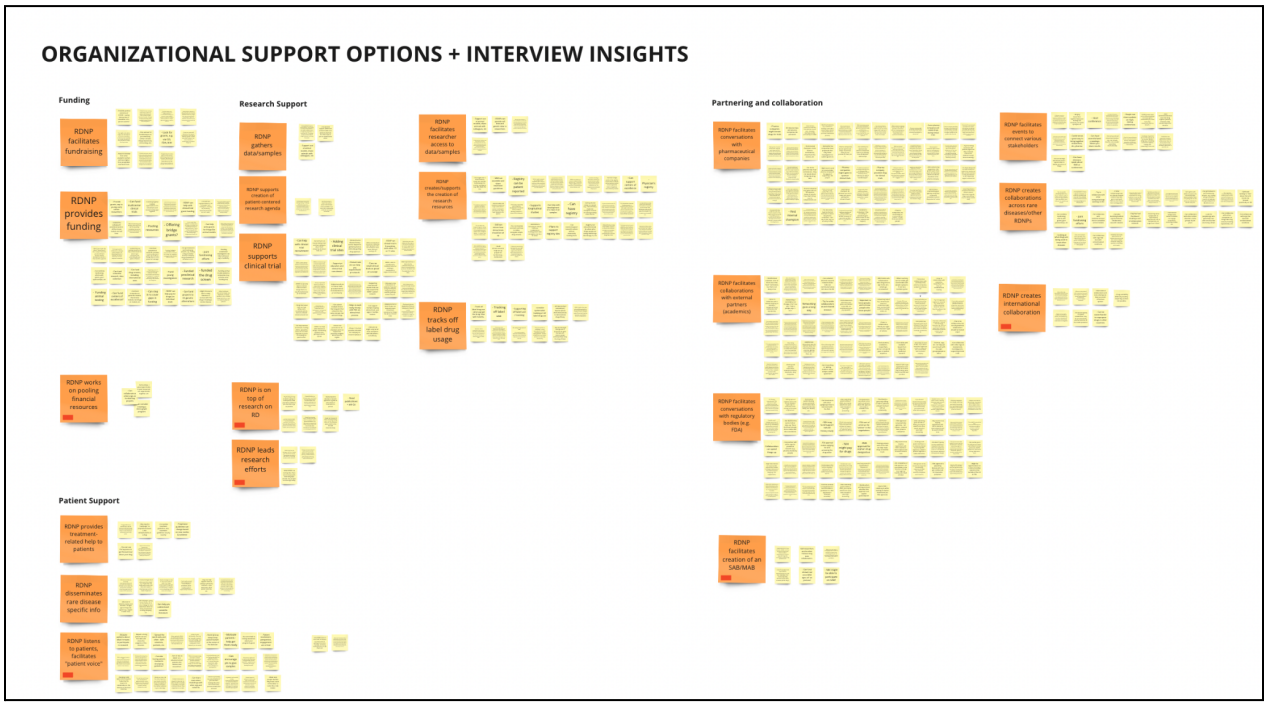


Figure 8: Process of theme and insight synthesis in MIRO

Synthesis & Tool Development

The qualitative insights from the interviews were then combined with the survey data and integrated into an interactive "ROADMAP" tool and data explorer interface. One of the most important aspects of this project is that it can help the rare disease community and that this data will be made available.

Thus, the ROADMAP project has produced several deliverables:

- The interactive data explorer
- The interactive ROADMAP tool
- Datasets
- Report
- Publication(s)

The interactive ROADMAP tool

The ROADMAP tool was developed using Jekyll, a static site generator that includes built-in hosting support for GitHub Pages. The decision to use Jekyll and GitHub pages was a tactical one: in addition to the user-friendly nature of both tools, we wanted to ensure that our tool – its code, documentation, etc. – was made available to users open-source, something that GitHub Pages readily supports. Within the Jekyll build environment, the structure of our tool was coded overwhelmingly with basic HTML/CSS/JavaScript, though we did occasionally use Liquid for the purpose of templating/automating. The content that populates our tool was written primarily using HTML, with the support of Twitter's Bootstrap 5 framework. The ROADMAP team is also appreciative of the Codepen community, whose open-source visualizations provided fruitful starting points for the development of various "proof of concept" iterations that were designed to test the use of JavaScript to support different interactivity features (e.g., scrolling, click-throughs, etc.).

OVERVIEW OF FINDINGS

The survey and our additional crowdsourced data provided a lot of unique insights from a variety of rare disease stakeholders. We are currently in the process of writing and submitting a publication with our research findings. Once this publication is available open-source, it will be available to read [HERE]. Until then, we provide preliminary, high-level findings below from each of the types of datasets we gathered as a part of this project. The data can be explored in our survey insights page here.

State of rare disease nonprofits in the US

From our crowdsourcing initiative, **711** rare disease nonprofit organizations were identified in the US and basic information was extracted from organizations' websites. Interestingly, 56% of these organizations were founded between 2011-2021. **430 (60.5%)** organizations focus on a single rare disease and its subtypes, while **264** focus on multiple rare diseases (some organizations were very broad, focusing on either all rare diseases or both rare

and common diseases). **416 (41.9%)** of the organizations were started by loved ones of patients (parents, spouses, siblings, etc.) and **206 (20.8%)** were started by the patients themselves.

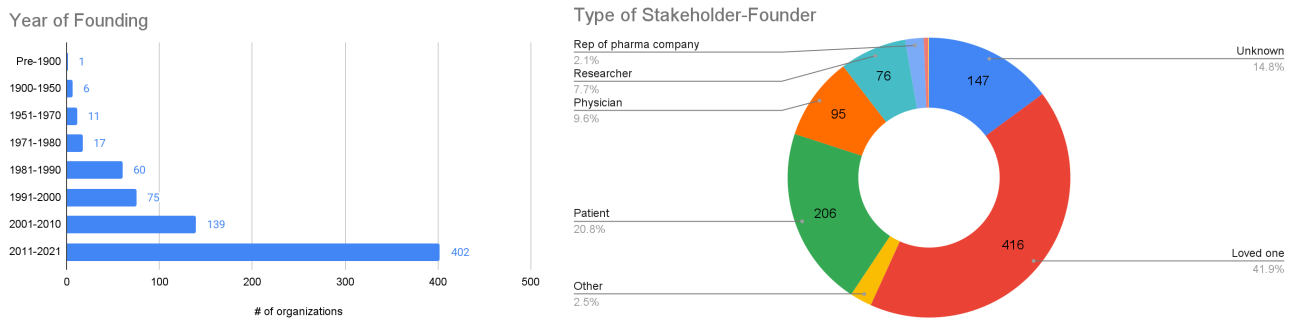


Figure 9: rare disease nonprofit year of founding (left) and type of stakeholder as the founder (right)

Utilizing the organization's websites, our team of extractors identified how many stated that they utilize certain resources that may be important for an organization to be able to successfully support research projects in the future, such as: **423 (59.5%)** organizations had a scientific advisory board, **335 (47%)** had a formalized research agenda, **291 (41%)** had a natural history study or registry, 2**34 (33%)** mentioned formalized treatment guidelines for their rare disease of focus, and **142 (20%)** had a biobank.

Interestingly for this project, drug repurposing initiatives, off-label use or treatment guidelines that included repurposed drugs were found on **135 (19%)** organizations' websites.

Off-Label Use according to an established medical resource

In parallel to the survey, we extracted data on the repurposing of drugs according to an established medical resource for a select group of 68 rare diseases, a subset from the ROADMAP survey data. The team of extractors extracted a total of **1,054** (**542** unique) drugs for these **68** rare diseases. **926** drugs were listed as recommended or potentially helpful, **53** were not recommended. **702 (71.7%)** were FDA approved for the rare disease in question, **203 (20.7%)** were being utilized off-label, and **74 (7.6%)** were not yet approved for anything in the US. **304 (25.7%)** unique drugs were recommended but either off-label or not yet approved for anything in the US, which illustrates just how common off-label drug use is and how huge the potential for drug repurposing is.

We also compared the two datasets: the data from this established medical resource and the data that met inclusion criteria in the ROADMAP survey by rare disease nonprofits. We found 33 rare diseases that were present in both datasets. Next, we looked at all the drugs (both off-label and FDA-approved) listed for each rare disease in both datasets.

In our sample, more than half 59/90 (65.5%) of the unique drugs listed for these 33 rare diseases in the ROADMAP survey by rare disease nonprofits are in data from the medical resource for the same rare diseases. The gap between what was mentioned in the survey and what was found in the medical resource could be explained by the fact that many of the drugs mentioned as "being repurposed" may not yet be used in patients or they may be referring to repositioned drugs that need an FDA-approval before they can be used in patients.

We also found that only 59/404 (14.6%) of the unique drugs listed in the medical resource were also reported in the survey by RDNPs. So, a large number of drugs recommended in the medical resource were not mentioned in the survey. This could represent drugs that were intentionally developed for the disease of interest and/or a lack of awareness of drugs that have been or are being repurposed for the disease of interest, among others.

Survey insights: Rare Disease Nonprofits

The state of drug repurposing projects, led by rare disease nonprofit organizations

Among the **147** organizations surveyed, **127** support research. **58** organizations are currently pursuing drug repurposing projects and another **58** are not yet, but are interested in doing so. Out of the **58** which are pursuing drug repurposing, only **40** organizations provided a specific drug of focus. This may indicate that the remaining 18 organizations are in preliminary stages and have not yet identified a specific drug yet. Out of the 40 that already have one or more drugs for repurposing, **14** have met at least one endpoint. There is a wide range in terms of what organizations selected as their success endpoint goal: "Drug to provide significant reduction in symptoms" (**48 selected** this endpoint - **17** reported that is has been met), "Drug to provide significant improvement in quality of life" (**48** selected - **11** met), and "Drug to be freely available to patients off-label with safety / efficacy data" (**47 selected** - **2** met).

While **5** drugs have made it to FDA approval, only **40 (43%)** of drug repurposing projects in our data even set FDA approval as their success endpoint goal. Thus, getting an FDA approval for the drug to be used in a new disease area is often less of a goal of interest than demonstrating it is effective and helping patients, which can be accomplished without an FDA approval (as long as it is approved for another disease and widely available).

Pursuit of multiple projects

Among the **58** organizations which reported to be currently pursuing or have pursued drug repurposing projects in the past, **19** organizations have pursued multiple drugs for repurposing, with an average of **3.84** projects per organization (range of 2-7); most organizations only pursued two projects at a time, likely due to the financial limitations of supporting multiple complex projects.

Off-label use tracking

"Off-label drug use" refers to when a drug is prescribed for a disease that it is not specifically approved for. Out of the **58** organizations that are pursuing drug repurposing projects (or have in the past), only **17** systematically track off-label drug use in their patient population. This is important to highlight since tracking off-label drug use is an important way to assess whether a repurposed drug is effective. Data on off-label drug use can also be used as a way to identify a drug for potential further repurposing research. In fact, data analysis of off-label drug use was the third most reported method for identifying a repurposed drug, helping **14** organizations to identify **25** drugs as promising for their rare diseases. Though the majority do not track off-label drug use at this time, **66**% stated that they are interested in tracking this information in the future.

Identifying a promising drug for repurposing

The most common method for identifying promising drug repurposing opportunities was preclinical/translational research (**66 drugs**). This category can also be described as pathogenesis targeting, whereby the researcher identifies a potential problem underlying a disease (e.g., mTOR activation is increased in Castleman disease) and

then matching a drug to reverse the problem (e.g., using an mTOR inhibitor to treat Castleman disease), which is then studied further in the laboratory. The other top choices were looking at drugs used in similar diseases (**29**), off-label use data analysis (**25**), high throughput drug screening (HTDS) (**21**) and literature review/meta-analysis (**16**). The least common option was machine learning/artificial intelligence approaches (**2**), which may speak to the novelty of this approach and its slow integration into the existing research/repurposing processes or that it is mostly being utilized without the involvement of rare disease nonprofit organizations. It is also important to mention that more than one identification method was employed to identify **57** drugs. The most common combinations of methods were: 1) **HTDS** and **Preclinical/Translational research** (**11** cases); and 2) **looking at drugs used in similar diseases** and **Preclinical/Translational** (**9** cases).

It is also interesting to point out that **11** drugs were identified as promising by multiple organizations using different drug identification methods. This speaks to the value of pursuing multiple avenues at once and triangulating the findings from one method with another.

Current stage of drugs being repurposed

Among the **94** (**76** unique) drugs which are in the process of being repurposed by **40** organizations, we categorized the drug repurposing project one of several stages:

- **Early stage:** Securing funding, trying to secure researcher interest to pursue, testing existing drugs in mouse or other animal models, patient data collection
- **Clinical trials:** Recruiting patients for clinical trials, executing clinical trials stages I, II, III
- **Late stage:** Analyzing clinical trial data, submitting for FDA approval, receiving FDA Breakthrough designation

In our dataset, **25** drugs are in early stages, **28** are in clinical trials, **14** are in late stages.

Defining success in drug repurposing

Additionally, we worked to develop several success outcome categories, which would include more than just official FDA approval and provide alternative ways to conceptualize and measure drug repurposing success for rare diseases.

Our resulting outcome categories include:

- **FDA approval**
- **Off-label use, with some subjective measure of benefit** (Drug to provide significant reduction in symptoms, improvement in quality of life, increase life expectancy/decrease in mortality, cure of disease, prevention of relapse)
- **Off-label use, without any subjective measure of benefit**
- **Unsuccessful (project abandoned or FDA approval denied)**

In our dataset, **36** drugs have reached an outcome (FDA approval / off-label use / unsuccessful) of interest. Only **5** drugs have so far made it to FDA approval for their new rare disease indication but **18** are being used off-label and having some sort of positive effect on the patients despite not having official FDA approval. **13** drugs were reported to have been abandoned. No drugs in our dataset fell into the third category, off-label use, without any

subjective measure of benefit. Most drugs (**58**) are still in process, and have not yet achieved any of the defined endpoint categories.

Note: due to some overlap between the stage and the outcome categories, the number of drugs in the stages does not add up exactly to the ones marked as in progress in the outcome

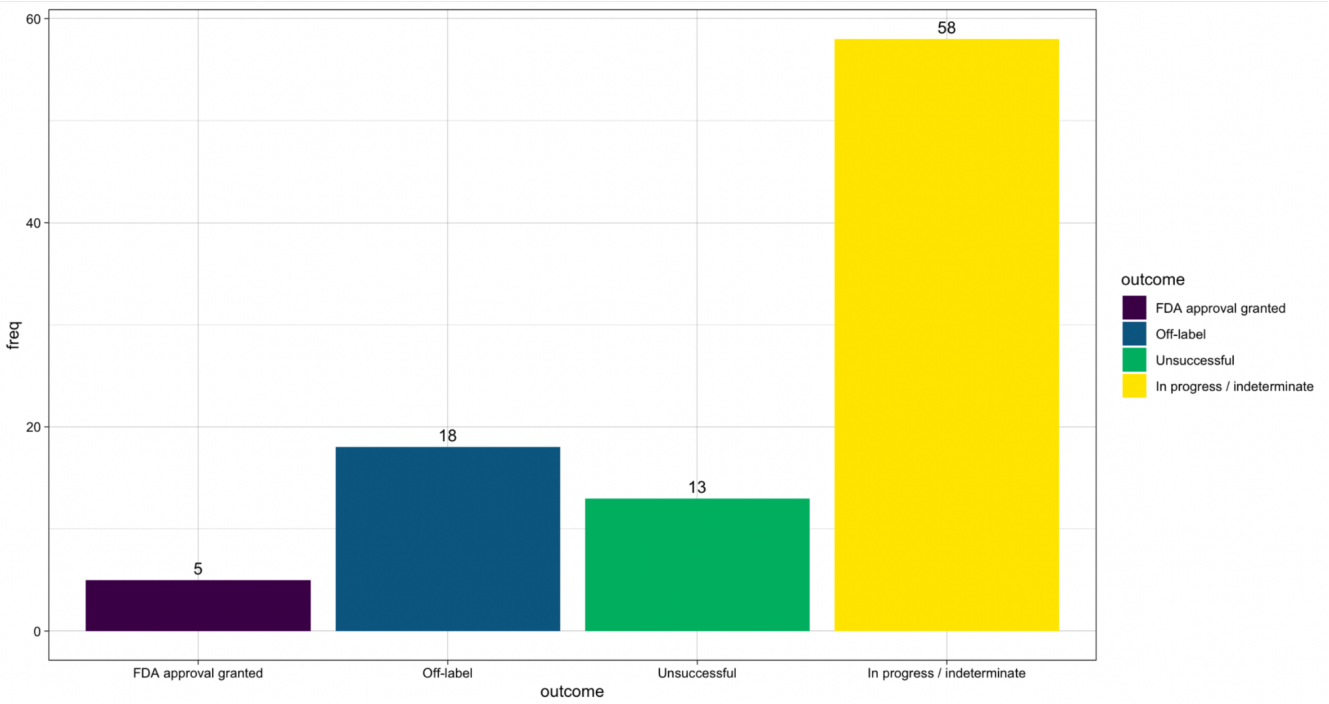


Figure 10: Categorization of various endpoints for drug repurposing projects

High-throughput drug screening

Though many people think of high-throughput drug screening (HTDS), the process whereby lots of drugs are tested on cells or other models from a given disease to see if anything brings about a change of interest, as the primary mechanism for identifying drug repurposing opportunities, it is clear from this survey that several other methods are used to identify repurposed treatments. In our data, **21** drugs were identified via high-throughput drug screens. Out of these drugs, **6** are still in early stages, **6** are in clinical trials, **5** are in late stages, and **1** has made it to FDA approval. **6** are being prescribed off-label, and **3** have been unsuccessful.

FDA-approval

Out of our **94** (**76** unique) drugs, only **5** have made it to FDA approval:

1. **Dupilumab** // Eosinophilic diseases
2. **Selumetinib** // Neurofibromatosis
3. **Alpelisib** // CLOVES syndrome
4. **Rituximab** // Pemphigus, Pemphigoid
5. **Sirolimus** // Lymphangioleiomyomatosis

Looking closer at the 5 organizations which have had a repurposed drug make it to FDA approval, they are on average **25.8** years old (range: **11** - **44**); their annual funding ranged from **$100,000** to **more than $5,000,000** (the most common selection was "**$1,000,000 and $2,000,000**"); and they have the following characteristics: **all 5** have an SAB/MAB, **3** have a natural history study, **3** have a formal research agenda, **3** already have an FDA approved drug prior to pursuing drug repurposing, **2** have a patient registry, and **1** has a biobank. One of these organizations has no full-time staff, relying entirely on volunteer or part-time staff to achieve their success, while the other four have anywhere from **1** to **40** full-time staff.

The amount of time that passed between initial FDA approval and repurposing approval ranged greatly, from 3 up to 21 years. Also worth mentioning that Selumetinib was a case of drug repositioning, as it was being explored as a potential drug for several indications, but was not pursued all the way to FDA approval; instead, it was identified as promising for an alternative use and received its first FDA approval for Neurofibromatosis.

It is important to note that there are very many factors that affect whether a drug can ever get FDA approval that do not depend on its safety, efficacy and are beyond the control of a researcher or a rare disease nonprofit organization supporting a drug repurposing project. Notably the most important of these is whether the pharmaceutical company who developed the drug is interested in investing the time and money into supporting the FDA approval application for a new disease when it is already approved for another disease. Furthermore, FDA approval for the new disease is not always necessary in order to have a drug be able to reach patients in need (the primary purpose). Thus, we consider FDA approval to be one of many metrics of success.

Off-label use with some subjective measure of benefit

As an alternative success outcome to FDA approval, we can consider off-label, with some subjective measure of benefit, such as being freely available to patients off label with safety / efficacy data, providing significant reduction in symptoms, improvement in quality of life, increase life expectancy / decrease in mortality, provide cure of disease, provide prevention of relapse. If we look at organizations that fit this criteria for at least one drug, we end up with **12** organizations. They are on average **17.6** years old (range: **2-44**); the majority (**4 organizations**, **33.33**%) reported annual funding between **$100,000 and $500,000;** and they have the following characteristics: **11** (**91.67**%) have an SAB, **6** (**50**%) have a natural history study, **7** (**58.33**%) have a formal research agenda, **6** (**50**%) have a patient registry, **5** (**41.67**%) have a biobank, and **3** (**25**%) already have an FDA-approved drug prior to pursuing drug repurposing (one organization has two FDA-approved drugs). Interestingly, **6** (**50**%) have no full-time staff, relying entirely on volunteer and/or part-time labor to achieve their success.

Among these **12** organizations, there are **44** unique drugs being repurposed. The most common drugs among these organizations were **Sirolimus** (**3** organizations), **Trametinib, Everolimus, and Bevacizumab** (**2** organizations each). The most common identification method for these drugs was **Preclinical/Translational research** (**30**), closely followed by **data from similar diseases** (**17**) and **off-label use** (**17**). Most of these drugs are currently in early stages or clinical trials, specifically in **recruiting patients for clinical trials** (**18**). Their respective rare diseases have the following characteristics: **11** (**91.67**%) have animal models, **8** (**66.67%**) have cell lines developed, **9** (**75%**) have an identified genetic mutation, **8** (**66.67%**) have an ICD code, **7** (**58.33%**) have treatment guidelines, **5** (**41.67**%) have a clear understanding of etiology or disease pathogenesis, and 3 (**25**%) have predictive biomarkers.

Correlational insights

We compared data on drug identification methods with the stages of the drug repurposing process and success outcomes. In the table below, drug repurposing stages are represented in the rows, while drug identification methods are represented in the columns.

	medical data (y)	medical data (n)	literature (y)	literature (n)	similar diseases (y)	similar diseases (n)	pre/trans (y)	pre/trans (n)	HTDS (y)	HTDS (n)	AI/ML (y)	AI/ML (n)	other (y)	other (n)
STAGE														
early (y)	6	19	2	22	2	23	14	11	6	19	1	24	6	19
early (n)	14	34	11	37	20	28	38	10	11	37	1	47	7	41
clinical (y)	6	21	6	21	9	18	24	3	6	21	1	26	4	23
clinical (n)	14	32	8	38	13	33	38	18	11	35	1	45	9	37
late (y)	6	8	3	11	6	8	9	5	5	9	0	14	3	11
late (n)	14	45	11	48	16	43	43	16	12	47	2	57	10	49
OUTCOME														
fda (y)	2	3	0	5	3	2	5	0	1	4	0	5	0	5
fda (n)	23	63	16	70	26	60	61	25	20	66	2	84	15	71
off-label (y)	10	9	4	15	9	10	14	5	6	13	0	19	0	19
off-label (n)	13	43	9	47	15	41	41	15	14	42	2	54	10	46
unsuccessful (y)	0	10	2	11	4	9	9	4	3	10	0	13	2	11
unsuccessful (n)	22	56	14	64	25	53	57	21	18	60	2	76	13	65
indeterminate (y)	11	46	10	47	16	41	41	16	12	45	2	55	13	44
indeterminate (n)	14	20	6	28	13	21	25	9	9	25	0	34	2	32

	medical data (y)	medical data (n)	literature (y)	literature (n)	similar diseases (y)	similar diseases (n)	pre/trans (y)	pre/trans (n)	HTDS (y)	HTDS (n)	AI/ML (y)	AI/ML (n)	other (y)	other (n)
STAGE														
early (y)	8.22%	26.03%	2.78%	30.56%	2.74%	31.51%	19.18%	15.07%	8.22%	26.03%	1.37%	32.88%	8.22%	26.03%
early (n)	19.18%	46.58%	15.28%	51.39%	27.40%	38.36%	52.05%	13.70%	15.07%	50.68%	1.37%	64.38%	9.59%	56.16%
clinical (y)	8.22%	28.77%	8.22%	28.77%	12.33%	24.66%	28.92%	3.61%	8.22%	28.77%	1.37%	35.62%	5.48%	31.51%
clinical (n)	19.18%	43.84%	10.96%	52.05%	17.81%	45.21%	45.78%	21.69%	15.07%	47.95%	1.37%	61.64%	12.33%	50.68%
late (y)	8.22%	10.96%	4.11%	15.07%	8.22%	10.96%	12.33%	6.85%	6.85%	12.33%	0.00%	19.18%	4.11%	15.07%
late (n)	19.18%	61.64%	15.07%	65.75%	21.92%	58.90%	58.90%	21.92%	16.44%	64.38%	2.74%	78.08%	13.70%	67.12%
OUTCOME														
fda (y)	2.20%	3.30%	0.00%	5.49%	3.30%	2.20%	5.49%	0.00%	1.10%	4.40%	0.00%	5.49%	0.00%	5.49%
fda (n)	25.27%	69.23%	17.58%	76.92%	28.57%	65.93%	67.03%	27.47%	21.98%	72.53%	2.20%	92.31%	16.48%	78.02%
off-label (y)	13.33%	12.00%	5.33%	20.00%	12.00%	13.33%	18.67%	6.67%	8.00%	17.33%	0.00%	25.33%	0.00%	25.33%
off-label (n)	17.33%	57.33%	12.00%	62.67%	20.00%	54.67%	54.67%	20.00%	18.67%	56.00%	2.67%	72.00%	13.33%	61.33%
unsuccessful (y)	0.00%	11.36%	2.20%	12.09%	4.40%	9.89%	9.89%	4.40%	3.30%	10.99%	0.00%	14.29%	2.20%	12.09%
unsuccessful (n)	25.00%	63.64%	15.38%	70.33%	27.47%	58.24%	62.64%	23.08%	19.78%	65.93%	2.20%	83.52%	14.29%	71.43%
indeterminate (y)	12.09%	50.55%	10.99%	51.65%	17.58%	45.05%	45.05%	17.58%	13.19%	49.45%	2.20%	60.44%	14.29%	48.35%
indeterminate (n)	15.38%	21.98%	6.59%	30.77%	14.29%	23.08%	27.47%	9.89%	9.89%	27.47%	0.00%	37.36%	2.20%	35.16%

Figure 11: Contingency table looking at drug identification methods and what stage the drug is currently in

From this table, we can see that preclinical/translational research seems to be the approach most often associated with clinical or late stages of the repurposing process, and in the outcomes, it is most associated with FDA approval or the alternative success outcome, off label use with measures of success.

We also correlated the organizational characteristics with the drug repurposing stages. In the table below, drug repurposing stages are represented in the rows, while organizational resources are represented in the columns.

	contact registry (y)	comm discussion (y)	comm gatherings (y)	comm navigator (y)	conferences (y)	fundraising events (y)	sab/mab (y)	pt registry (y)	natural history (y)	PRO (y)	biobank (y)	research agenda (y)
STAGE												
early (n)	15	19	18	12	15	19	19	10	10	8	8	11
early (y)	11	13	8	10	10	9	12	9	8	4	4	9
clinical (n)	10	13	11	8	10	10	11	8	7	5	3	7
clinical (y)	16	19	15	14	15	18	20	11	11	7	9	13
late (n)	19	23	18	16	19	19	22	15	14	9	11	16
late (y)	7	9	8	6	6	9	9	4	4	3	1	4
OUTCOME												
fda (n)	29	34	26	22	26	29	33	21	22	13	15	23
fda (y)	4	5	5	4	5	5	5	2	3	2	1	3
off-label (n)	18	20	15	12	16	15	19	11	12	5	7	12
off-label (y)	9	13	11	9	10	13	13	7	8	5	5	9
unsuccessful (n)	23	29	23	18	23	25	27	14	17	9	10	17
unsuccessful (y)	10	10	8	8	8	9	11	9	8	6	6	9
indeterminate (n)	0	0	0	0	0	0	0	0	0	0	0	0
indeterminate (y)	18	28	23	20	22	23	27	18	16	11	10	18

	contact registry (y)	comm discussion (y)	comm gatherings (y)	comm navigator (y)	conferences (y)	fundraising events (y)	sab/mab (y)	pt registry (y)	natural history (y)	PRO (y)	biobank (y)	research agenda (y)
STAGE												
early (n)	45.45%	57.58%	54.55%	37.50%	46.88%	57.58%	57.58%	31.25%	31.25%	25.81%	25.00%	33.33%
early (y)	33.33%	39.39%	24.24%	31.25%	31.25%	27.27%	36.36%	28.13%	25.00%	12.90%	12.50%	27.27%
clinical (n)	30.30%	39.39%	33.33%	25.00%	31.25%	30.30%	33.33%	25.00%	21.88%	16.13%	9.38%	21.21%
clinical (y)	48.48%	57.58%	45.45%	43.75%	46.88%	54.55%	60.61%	34.38%	34.38%	22.58%	28.13%	39.39%
late (n)	57.58%	69.70%	54.55%	50.00%	59.38%	57.58%	66.67%	46.88%	43.75%	29.03%	34.38%	48.48%
late (y)	21.21%	27.27%	24.24%	18.75%	18.75%	27.27%	27.27%	12.50%	12.50%	9.68%	3.13%	12.12%
OUTCOME												
fda (n)	72.50%	85.00%	66.67%	57.89%	66.67%	72.50%	82.50%	53.85%	56.41%	34.21%	38.46%	57.50%
fda (y)	10.00%	12.50%	12.82%	10.53%	12.82%	12.50%	12.50%	5.13%	7.69%	5.26%	2.56%	7.50%
off-label (n)	52.94%	58.82%	44.12%	36.36%	48.48%	44.12%	55.88%	33.33%	36.36%	15.63%	21.21%	35.29%
off-label (y)	26.47%	38.24%	32.35%	27.27%	30.30%	38.24%	38.24%	21.21%	24.24%	15.63%	15.15%	26.47%
unsuccessful (n)	57.50%	72.50%	58.97%	47.37%	58.97%	62.50%	67.50%	35.90%	43.59%	23.68%	25.64%	42.50%
unsuccessful (y)	25.00%	25.00%	20.51%	21.05%	20.51%	22.50%	27.50%	23.08%	20.51%	15.79%	15.38%	22.50%
indeterminate (n)	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
indeterminate (y)	64.29%	100.00%	82.14%	74.07%	81.48%	82.14%	96.43%	66.67%	59.26%	42.31%	37.04%	64.29%

Figure 12: Contingency table looking organizational characteristics and where the drug is currently

From this table, we can see that having an SAB and community discussion platform appear to be the factors most associated with the clinical phase of repurposing and not being unsuccessful.

Collaboration across rare diseases

It is very interesting that the second most-selected drug identification method was **looking at drugs used in similar diseases**, as one of our research focuses was the levels of collaboration between different organizations, and we were specifically interested to see how much collaboration we would see among different rare disease nonprofits. With the data from the collaboration section of the survey (120 out of the total 147 of nonprofits provided this info), we were able to build networks. Each node in a network represents an organization, and each edge signifying a connection between them, either an explicitly mentioned connection via the survey data or an affiliation edge which was later added by the research team. The nodes are colored by type of organization, with a distinction made between our participating rare disease nonprofit organizations (Rare Disease Non Profit (RDNP), and other nonprofit organizations mentioned by the respondents (Non profit):

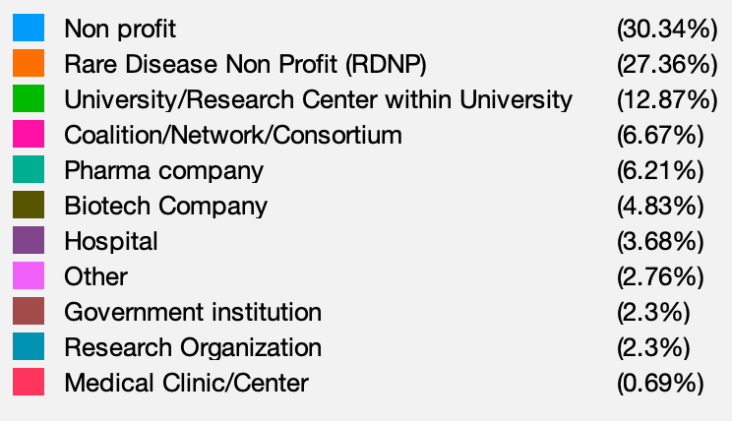


Figure 13: Network node color legend

From looking at the network visualization, we can see that most of the rare disease nonprofit organizations in our data are actually part of one large "connected component". A connected component means that there is a path between every pair of nodes. In other words, it is a group of nodes that are all reachable from one another by following the edges of the graph. The nodes here are sized by degree, with the largest two nodes being NORD and Global Genes, which is not surprising since these are large, umbrella organizations that each have hundreds of member organizations.

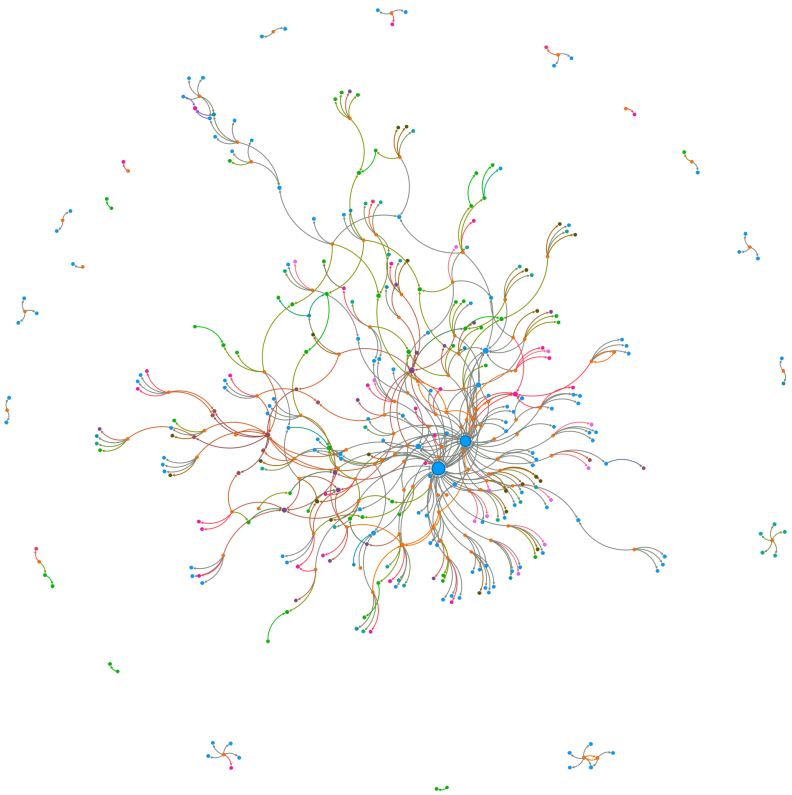


Figure 14: Collaboration network from self-reported "top 5" collaborators from our ROADMAP survey data

Additionally, since we asked organizations about the types of collaborations they are engaged in with their top 5 closest collaborators, we are able to view subsets of the networks based on the type of activity. For example, we

can see organizations who share data, share resources and share knowledge about their experiences (See Figure 12, in which green edges = yes, red edges = no). There is some overlap within those who come up as doing all three (consistently green edges across all three networks below), but there are some significant differences. For one, knowledge sharing seems to be a lot more common, as compared to data or resource sharing (evident by the fact that there are a lot more green edges, compared to the other two networks). Also, even though NORD and Global Genes are the highest nodes by their number of connections to other organizations overall, they were not reported to be in the top 5 organizations with which our participating nonprofits engage with data sharing or resource sharing with (evident by the fact that most of the edges in these two networks are red). This may indicate that even though these types of data sharing might still be happening, rare disease nonprofits work more closely with other nonprofit or academic institutions for these types of collaborations.

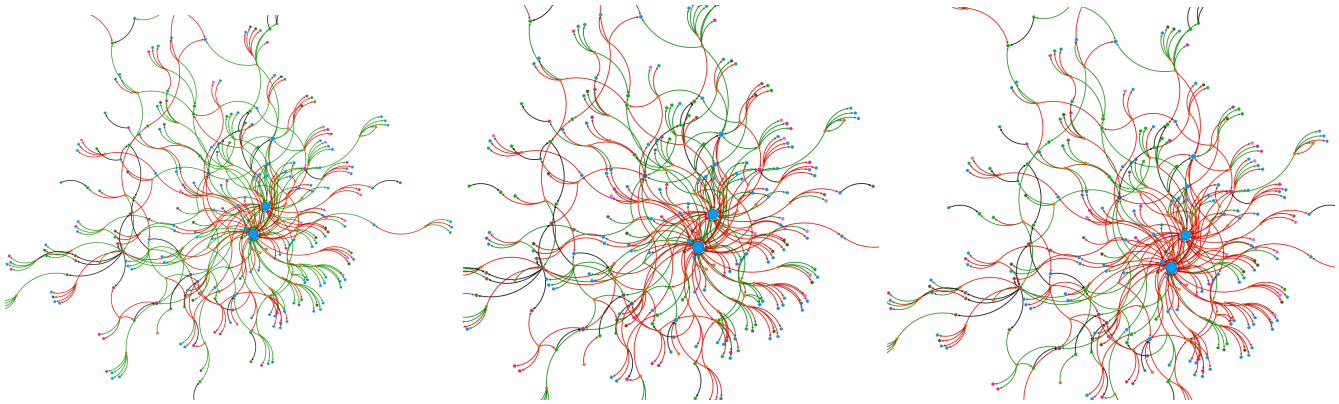


Figure 15: Left to Right: Knowledge sharing, Resource sharing and Data sharing, where green edges = yes, red edges = no

You can explore more of the rare disease nonprofit data on the ROADMAP survey data explorer.

Survey insights: Patients & Loved ones

For rare disease patients and their loved ones (these stakeholders received the same set of questions), we were interested in their experience with off-label drugs for their rare disease, including roadblocks in access, as well as their general attitude towards and comfort level with drug repurposing.

While the majority of patients are aware of whether the drugs they are taking are FDA-approved or off-label, **28.8%** stated they are not aware of the status of the drug that they are on.

Most patients (**76.6%)** have heard of the term “off-label” but have not talked to their doctor about using an off-label drug, nor are most of these patients (**66.5%**) aware of other patients using off-label drugs. Generally, patients are **very willing** to follow their doctor's recommendations based on **trust with their doctor,** as well as **drug safety and efficacy**. **109** patients received the survey from one of the 58 rare disease nonprofits which stated that they are currently or were previously involved in drug repurposing; **43 (39.4%)** of these **109** patients said they were not aware of any drug repurposing initiatives being done for their rare disease. Of note, these

results contrasted with the large proportion of off-label use reported by rare disease nonprofit organizations and physicians (see below).

Patients stated that they were **extremely interested (77/296, 26%)** in drug repurposing with the most impactful aspect being having more treatment options available **(232/296, 78.4%).** Respondents strongly agreed that drug repurposing is **generally safe (32/296, 10.8%)** and **necessary (118/296, 40%)** for their rare disease. Patients reported no roadblocks in access to **301** drugs **(65%)** that they have been prescribed**.** This insight was interesting given that anecdotal reports suggest that insurance not approving off-label drugs can greatly affect patient access to potentially life-saving treatment, even if the drug repurposing process has been successful.

Overall, it seems that the patients and loved ones who participated in the survey have a positive attitude towards drug repurposing, but some are now aware of ongoing repurposing efforts or off-label drugs that may be promising for their rare disease. These insights can be useful for rare disease nonprofits to be aware of in crafting their communication materials, to make sure that their patient populations are aware of all the research being pursued and all treatment options available with sufficient safety and efficacy data, even if they are not yet FDA-approved for their rare disease.

You can explore the patient/loved one data more on the ROADMAP survey data explorer.

Survey insights: Physicians

For physicians, we were mainly interested in the frequency and factors taken into consideration when deciding to prescribe a drug off-label, off-label drug prescription practices for the rare diseases they treat,and their levels of involvement in clinical trials for rare diseases.

Due to the small sample of physicians who participated in the survey (**21**), it is difficult to make broad generalizations. Among the physicians in our dataset, **70.6%** of them were aware of off-label drug options for the rare diseases that they treat and they prescribe them often.

A little less than half of the physicians who took the survey **(10/21, 47.62%**) said that they prescribe specific off-label treatments to their rare disease patients; these physicians listed a combined 26 drugs. When asked how effective these treatments are at reducing overall symptoms, physicians described **18/26 (69.23%)** of the drugs that they listed as “somewhat effective.” Physicians described **17/26 (65.38%)** drugs as “somewhat effective” at preventing worsening of their patients’ rare disease, and **20/26 (76.92%)** as “somewhat effective” at improving their patients’ overall state of health.

Physicians in our data reported that **80%** of the current off-label options do not completely address all of the challenges associated with their patients’ rare disease.

More than half of physicians reported being at least “somewhat” familiar with drug repurposing **(12/17, 70.59%)** and being aware of ongoing clinical trials (**11/17, 64.71%**). Of those aware of ongoing clinical trials, all but two physicians **(9/11, 81.82%**) said that they have referred patients of theirs to these trials. **100%** of our participating

physicians stated that, if the opportunity presented itself, they would be willing to be a co-investigator on a clinical trial to assess the efficacy of a drug that is being repurposed for a rare disease that they treat.

Although the sample size of participating physicians was small (21), we present the data as an initial reference for understanding physician involvement in different aspects of drug repurposing initiatives. Making broad generalizations as to physicians that treat rare disease based on this data may be challenging based on the small sample size. You can explore the physician data on the ROADMAP survey data explorer.

Survey insights: Researchers

For researchers, we asked questions regarding their involvement in research projects to advance drug repurposing for any rare diseases, what their motivations are and what challenges they've encountered.

We had **43** researchers participate in the ROADMAP survey, and most of them stated that they are **extremely interested (20/43, 46.51%)** in drug repurposing and find it **extremely important (22/43, 51.16%)**.

Among the researchers who reported that their drug repurposing research is ongoing, 14 reported to be working on understanding disease mechanisms, 12 are searching for drug targets. 25 researchers reported that there have been drugs identified; among these, 10 are in the process of being considered which patients they are most likely to help, 8 have some preclinical data generated, 7 have not had validated in preclinical data. Among the drugs that the researchers are focused on, **sirolimus** was the highest reported drug (**5** cases), with **ruxolitinib** coming in second (**3** cases).

Among the **31 (97%)** of researchers that have done research related to drug repurposing in the past or are currently, the top motivation is its **impact on patients**. Their top difficulties were **lack of funding (20/43, 46.51%), lack of support/collaboration from pharmaceutical companies (13/43, 30.23%),** and **lack of biospecimens (8/43, 18.6%)**.

Importantly, most **(30, 83.3%)** of researchers in our data reported engaging the rare disease patient population at least in some stage of their research process. Patients bring unique perspectives and insights into their conditions, which can inform research priorities and improve the relevance and quality of research outcomes.

Among those researchers that reported to be directly involved in the dissemination of this information to the rare disease community, most **(8, 80%)** reported accomplishing this task by working with a rare disease nonprofit organization, though social media and traditional media outlets are also utilized to a lesser extent. Although the sample size of participating researchers was small (43), we present the data as an initial reference for understanding their involvement in different aspects of drug repurposing initiatives. Making broad generalizations as to researchers whose research relates to rare disease drug repurposing based on this data may be challenging based on the small sample size.

You can explore the researcher data more on the ROADMAP survey data explorer.

ACKNOWLEDGEMENTS

We wish to thank the following individuals and organizations for their contributions to this important work:

CDCN Team

Dr. David Fajgenbaum, Ania Korsunska, Mileva Repasky, Mary Zuccato, Sarah E. Bolden, Trae Boyd and prior CDCN team members Johnson Khor and Dr. Ruth-Anne Langan Pai.

Participating rare disease nonprofit organizations

Acromegaly Community Inc, Adult Polyglucosan Body Disease Research Foundation, Aicardi-Goutieres Syndrome Americas Association, AKU Society of North America, Alport Syndrome Foundation, Amyloidosis Support Groups Inc., Angelman Syndrome Foundation, ASXL Rare Research Endowment Foundation, Autoinflammatory Alliance, Barth Syndrome Foundation, Ben's Friends Patient Communities, Beyond Batten Disease Foundation, Bow Foundation, BPAN Warriors, CACNA1A Foundation, Campaign Urging Research for Eosinophilic Disease, Cardio Facio Cutaneous International, Castleman Disease Collaborative Network, CDG CARE, Champ Foundation, CHAMP1 Research Foundation, Chelsea's Hope Lafora Children's Research Fund, Child Neurology Foundation, Children's Tumor Foundation, Cholangiocarcinoma Foundation, Chordoma Foundation, Choroideremia Research Foundation, Chronic Recurrent Multifocal Osteomyelitis Foundation, Circadian Sleep Disorders Network, CLOVES Syndrome Community, Coalition to Cure Calpain 3, Congenital Hyperinsulinism International, Costello Syndrome Family Network (CSFN), CSNK2A1 Foundation, Cure CMD, Cure GM1 Foundation, Cure HHT, Cure JM Foundation, Cure VCP Disease, Inc., cureCADASIL Association, CureSearch for Children's Cancer, Cyclic Vomiting Syndrome Association, Daphne's Lamp, DDX3X Foundation, DeSanto-Shinawi Syndrome, Dreamsickle Kids Foundation, DRESS Syndrome Foundation, Dup15q Alliance, E.WE Foundation, Emily's Entourage, Epilepsy Alliance America, Facial Pain Association, Familial Chylomicronemia Syndrome (FCS) Foundation, Fanconi Anemia Research Fund, Foundation for CAMK2 Therapeutics, Foundation for Ichthyosis and Related Skin Types, Foundation for Sarcoidosis Research, FOXG1 Research Foundation, FPIES Foundation, GACI Global, Glut1 Deficiency Foundation, Hannah's Hope Fund, Hermansky-Pudlak Syndrome Network, Hugs For Mito, Inc., Huntington's Disease Society of America, Hypersomnia Foundation, Hypoparathyroidism Association, IDefine, Indian Organization for Rare Diseases, Infantile Neuroaxonal Dystrophy Cure Foundation, International Autoimmune Encephalitis Society, International Fibrodysplasia Ossificans Progressiva Association, International Pemphigus Pemphigoid Foundation, Jamal's Helping Hands Inc., Jansen de Vries Syndrome Foundation, Jansen's Foundation, Jordan's Guardian Angels, Kabuki Syndrome Foundation, KIF1A.ORG, Koolen-de Vries Syndrome Foundation, LAM Foundation, Life Raft Group, Liv4TheCure, Lymphangiomatosis & Gorham's Disease Alliance, Malan Syndrome Foundation, Mast Cell Hope, MEPAN Foundation, Mission: Cure, Mississippi Metabolics Foundation, Muscular Dystrophy Association, Myasthenia Gravis Foundation of America, Myositis Support and Understanding Association, National Foundation for Ectodermal Dysplasias, National Neutropenia Network, National Organization for Disorders of the Corpus Callosum, National PKU News, NEC Society, NephCure Kidney International, Neurodegeneration with Brain Iron Accumulation Disorders Association, Neurofibromatosis Northeast, Ocular Melanoma Foundation, Oligo Nation, Organic Acidemia Association, Pachyonychia Congenita Project, PCDH19 Alliance, Pityriasis Rubra Pilaris (PRP) Alliance, Platelet Disorder Support Association, Primary Ciliary Dyskinesia (PCD) Foundation, Project ALS, Progressive Familial Intrahepatic Cholestasis (PFIC) Network, PSC Partners Seeking a Cure, PTEN Hamartoma Tumor Syndrome Foundation, RASopathies Network, Raymond A. Wood Foundation, Recurrent Respiratory Papillomatosis Foundation, Remember The Girls, RUNX1 Research Program, RYR-1

Foundation, Sara's Cure AKA Clear Cell Sarcoma Foundation, SATB2 Gene Foundation, Shwachman-Diamond Syndrome Foundation, Siegel Rare Neuroimmune Association, SLC6A1 Connect, Smith-Kingsmore Syndrome Foundation, Smith-Lemli-Opitz /RSH Foundation, Snyder-Robinson Foundation, Sophie's Hope Foundation, Spastic Paraplegia Foundation, Spinal CSF Leak Foundation, Stiff Person Syndrome Research Foundation, STXBP1 Foundation, Sumaira Foundation, Superficial Siderosis Research Alliance, Syngap1 Foundation, T.E.A.M. 4 Travis (Together Ending Asplenia Mortality), Tatton Brown Rahman Syndrome Community, TBCK Foundation, Team Telomere, United States Thrombotic Microangiopathy Alliance, Usher 1F Collaborative, Usher Syndrome Coalition, Vici Syndrome Foundation, Inc, Pulmonary Alveolar Proteinosis (PAP) Foundation, TESS Research Foundation, and all their participating patients, loved ones, researchers and physicians.

Volunteers

Meg Zuccato, Anna Nguyen, Derek Ansel, Martin Lukac, Annalise Jear, Mitav Nayak, Panchatapa Baul, Samatha Hood, Sabina Grigorian, Jacob Lowy, Justin Wong, Sydney Grisham, Yuan (Abby) Feng, Leanna Chen, Megan Shieh, Bryan Aguilar, Marcy Spiker, Rose Weathers, Katherine Fang, Veikko Toikka, Robert Parillo, Lindsay McBride, Sara Barrett, Robert Parillo, Dallas Ryan, Veikko Toikka, Benita Balogun, Penny Deremer, Miti Patel, Emma Roemer, Neda Pazuki, Erikka Chowdhury, Stephanie Hage, Matt Scott, Geetha Turlapati, Jessica Xiang, Jada Watkins, Jazmin Loughlin, Justin Crawmer, Ferzana Niazi, Anaheit Arathoon, Michael Zhang, Vee Suresh, Mahima Sangtani, Sara Cronin, Rita Aberbach, Kayleigh Nicole Murray, Diane Baynes, Andrew Zhu, Anastasia Kakurina, Susanna Hunanyan, Jade Bondy, Simarsukh Dhillon, Angela Perry, Carolyn Canterbury, Owen Yu and Kanan Lathia.

Partners & Collaborators

Chan Zuckerberg Initiative, Medidata Solutions

Other

Beacon, Rare Revolution Magazine, CDRC, PTEN Research

Online resources

Codepen, Icons8, Bootstrap 5, MDB 5

REFERENCES

1. Haendel M, Vasilevsky N, Unni D, Bologa C, Harris N, Rehm H, et al. How many rare diseases are there? Nat Rev Drug Discov. 2020 Feb;19(2):77–8.
2. Orphan Drug Act [Internet]. 96 STAT. 2049, 97–414 Jan 4, 1983. Available from: https://www.fda.gov/media/99546/download
3. Kaufmann P, Pariser AR, 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 Journal of Rare Diseases. 2018 Nov 6;13(1):196.
4. Stone H. CURE ID: Capturing Clinician's Experiences Repurposing Drugs to Inform Future Studies in the Eraof COVID-19 [Internet]. 2020 Jun 9 [cited 2021 Apr 4]; FDA. Available from: https://www.fda.gov/about-fda/fda-pharmacy-student-experiential-program/fda-drug-topics-cure-id-capturing-clinicians-experiences-repurposing-drugs-inform-future-studies-era
5. Ghofrani HA, Osterloh IH, Grimminger F. Sildenafil: from angina to erectile dysfunction to pulmonary hypertension and beyond. Nature Reviews Drug Discovery. 2006 Aug;5(8):689–702.
6. Latif T, Chauhan N, Khan R, Moran A, Usmani SZ. Thalidomide and its analogues in the treatment of Multiple Myeloma. Experimental Hematology & Oncology. 2012 Sep 11;1(1):27.
7. Fajgenbaum D, Langan RA, Japp AS, Partridge HL, Pierson SK, Singh A, et al. Identifying and targeting pathogenic PI3K/AKT/mTOR signaling in IL-6 blockade–refractory idiopathic multicentric Castleman disease. J Clin Invest. 2019;129(10):4451–63.
8. RStudio Team. RStudio: Integrated Development Environment for R [Internet]. Boston, MA: RStudio, PBC.; 2020. Available from: http://www.rstudio.com/
9. Wickham H, Averick M, Bryan J, Chang W, McGowan LD, François R, et al. Welcome to the tidyverse. Journal of Open Source Software. 2019;4(43):1686.
10. Wickham H. stringr: Simple, Consistent Wrappers for Common String Operations [Internet]. 2022. Available from: https://CRAN.R-project.org/package=stringr
11. Wickham H, François R, Henry L, Müller K. dplyr: A Grammar of Data Manipulation [Internet]. 2022. Available from: https://CRAN.R-project.org/package=dplyr
12. Bryan J. googlesheets4: Access Google Sheets using the Sheets API V4 [Internet]. 2022. Available from: https://CRAN.R-project.org/package=googlesheets4
13. Bastian M, Heymann S, Jacomy M. Gephi: an open source software for exploring and manipulating networks. In: Third international AAAI conference on weblogs and social media. 2009.
14. Braun V, Clarke V. Using thematic analysis in psychology. Qualitative Research in Psychology. 2006 Jan;3(2):77–101.