# Landscape Review on Rare Disease Research Registries

# Final Report

Prepared For PCORI By: Melissa Adams, Nancy Berkman
Ellen Bishop, Kathleen Lohr, Huaqin Pan, Daniel Ringer
Meera Viswanathan, Nedra Whitehead

RTI International 3040 E. Cornwallis Road
Research Triangle Park, NC 27709

**May 2015**

This report was prepared under PCORI RFQ PCO-LANDRVWRAREDISEASE2014

## Acknowledgments

Emma Djabali, Danielle Whicher, Kara Odom Walker, David Hickam, Naomi Aronson, Sharon Terry, Yaffa Rubinstein, Jim Wu, Liz Horn, and Holly Peay for their input and review of this report.



Questions or comments may be sent to PCORI at info@pcori.org or
by mail to 1828 L Street NW, Suite 900, Washington, DC 20036.

## Overview

The Patient-Centered Outcomes Research Institute (PCORI) commissioned this landscape review to obtain a compilation of existing standards for rare disease (RD) research. The PCORI Advisory Panel on RD will use this landscape review to determine which existing standards PCORI should endorse or to provide insight on existing gaps. Looking at national and international standards, and with direction from the Panel's Working Group, we present information about best practices for developing RD registries and considering biospecimens and biobanks as a related activity. We also examine issues about study designs and how to evaluate the strength of evidence from RD research.

RDs are challenging for the patients who live with them, the physicians who diagnose and treat them, and the researchers who study them. We reviewed three approaches to addressing the challenges in studying RDs: registries, which make patients easier to locate and recruit, and provide efficient collection of standard data for analyses and monitoring; biobanks, which allow investigation of biomarkers without primary recruitment of patients; and study designs that are optimal for studies of the effectiveness of RD therapies.

Over the past several years, major advances have been made in developing RD patient registries and conducting RD research. Although inadequate data standardization and harmonization continues to present challenges to linking data across registries, new open-source registry platforms and common data elements provide the infrastructure needed to allow greater standardization. The development of virtual biobanks and of best practices for the management and governance of physical biobanks have increased the value of even small collections or small samples of biospecimens. New methodological research has resulted in study designs tailored for RDs or small populations, and reaching valid conclusions based on small bodies of evidence.

Our review did identify several areas that need further research. Most pressing may be the need to integrate policies and procedures for RD registries with best practices about designing and conducting studies and study design and grading strength of evidence. Similarly, identifying the types of analyses needed to answer important research questions and selecting the most robust and defensible methods are also critically important. Methodological research is needed to develop improved methods for the evaluation of the representativeness of RD registries that solicit participation by appeals on the Internet or from advocacy groups, and to investigate the validity of using registries to evaluate side effects and effectiveness of therapeutics after their approval for clinical use.

## Acronyms

| Acronym | Full Term |
|---|---|
| AHRQ | Agency for Healthcare Research and Quality |
| ALS | amyotrophic lateral sclerosis |
| CDE | common data element |
| CDM | common data model |
| CLIA | Clinical Laboratory Improvement Amendments |
| CORD | Canadian Organization for Rare Disorders |
| DBMD | Duchenne and Becker muscular dystrophies |
| EHR | electronic health record |
| eMERGE | electronic MEdical Records and GEnomics |
| EPC | Evidence-based Practice Center |
| EURODIS | (National Organization for Rare Disorders (NORD), Rare Diseases Europe |
| FDA | U.S. Food and Drug Administration |
| GRADE | Grading of Recommendations Assessment, Development, and Evaluation |
| GRDR® | Global Rare Diseases Patient Registry |
| HIPAA | Health Insurance Portability and Accountability Act |
| IC IRB | informed consent |
| NIH | Institutional Review Board |
| NINDS | National Institutes of Health<br>National Institute of Neurological Disorders and Stroke |
| NORD | National Organization for Rare Disorders |
| RCT | randomized controlled trial |
| RD | rare disease |
| SOE | strength of evidence |

# Contents

## Appendices



## List of Tables

# Introduction

The Patient-Centered Outcomes Research Institute (PCORI) Advisory Panel on Rare Disease (RD) commissioned this landscape review to provide guidance for RD research. Looking at national and international standards, and with direction from the Panel's Working Group, we present information about best practices for developing RD registries and considering biospecimens and biobanks as a related activity. We also examine issues about study designs and how to evaluate the strength of evidence from RD research.

Per PCORI's specifications, we focused on accepted and preferred ways to design new registries for research involving patients with RDs of all types. We did not conduct a formal systematic review for RD registries. Likewise, we focused on standards for study designs to use for future investigations of the efficacy or effectiveness of RD therapies and on methods for grading the strength of evidence from such trials or other studies. We did not conduct a systematic review purely of study designs or strength of evidence assessments for RD research.

The report has three main sections that address issues related to RD Research:

1. RD Research Registries
    a. Requirements of Registries for RD
    b. Management of Registries for RD
2. Stewardship of Biospecimens and Biobanks
3. Issues about Study Design and Strength of Evidence for RD Research

**Appendix A** documents the methods we applied to search the literature about RD registries and about questions relating to stewardship of biospecimens and biobanks.

**Appendix B** shows the methods we used to answer the following key questions for issues about study design and strength of evidence (Part 3):

1. What study designs can be used to evaluate therapies for patients with RDs? What are their applications and constraints?
2. What strength of evidence systems can be used when evaluating therapies for RD patients?

Because this report is a landscape review and not a systematic review, we included only those publications that we judged to be highly relevant or very recent.

## Burden on society

The definition of a rare disease (RD) varies considerably. The Rare Diseases Act of 2002 designates a RD as a condition that affects fewer than 200,000 people in the United States (*Rare Diseases Act of 2002* 2002). Other authorities define RDs differently, resulting in a range of the maximum prevalence from 1 to 6.3 affected per 10,000 people (Table 1). The Office of Rare Diseases of the United States National Institutes of Health (NIH) states that approximately 7,000 RDs have been identified (2015).

*Table 1. Definition and Maximum Prevalence of Rare Diseases by Country or Region*

| Country/region | Equivalent Maximum Prevalence per 10,000 people | Definition |
|---|---|---|
| Australia | 1 | A disease or condition likely to affect ≤ 2,000 individuals in Australia at any time (Australian Government 2014) |
| China | - | Conditions affecting <1/500,000 people or 1/10,000 neonates (Cui and Han 2015) |
| European Union | 5 | Life-threatening or chronically debilitating and prevalence < 1/2,000 people (EUR-Lex) |
| Japan | 4 | Conditions affecting < 50,000 people (Song et al. 2012) |
| South Korea | 4 | Diseases which affect ≤ 20,000 people that do not have appropriate treatment (Song et al. 2012) |
| Taiwan | 1 | Diseases which affect ≤ 10,000 people (Song et al. 2012) |
| United States | 6.3[1] | Conditions affecting < 200,000 people in the United States (*Rare Diseases Act of 2002* 2002) |

[1] For 2014.

Because RDs are rare and usually widely dispersed, they can be difficult to study. The underlying population may need to be extremely large to include the number of affected individuals needed for the research. Recruiting a small number of eligible individuals from a large population can be extremely difficult and expensive. Logistic challenges and resources often cause researchers to conduct studies with small convenience samples recruited from specialty clinics or facilities, which may not be representative of the population of individuals with the disease.

## Options for addressing the issue

We reviewed three approaches to addressing the issues of studying RDs: Part 1 research registries; Part 2 stewardship of biospecimens and biobanks; and Part 3 issues about study designs and strength of evidence.

## Part 1 Research Registries

Patient registries can make research possible that could not be conducted using other recruitment methods. A common definition of a patient registry is "…an organized system that uses observational study methods to collect uniform data (clinical and other) to evaluate specified outcomes for a population defined by a particular disease, condition, or exposure, and that serves one or more predetermined scientific, clinical, or policy purposes" (Gliklich, Dreyer, and Leavy 2014). Research registries require additional functionality, including "…storage, retrieval and dissemination…of data collected on identifiable individuals…" (Richesson and Vehik 2010). Including identifying information enables registries to follow individuals over time (Brooke 1974); to link the registry data to other data sources, such as clinical records or vital records; to link to other registries; to identify duplicate records; and to distinguish between relatives or individuals with the same first and last names.

Ample guidance has been written about creating and conducting patient registries for diseases or conditions that are not rare (Brooke 1974; Gliklich, Dreyer, and Leavy 2014; Gliklich et al. 2012). Although much of this guidance also applies to registries for RDs, the small number of affected individuals and the complexity of many RDs require special considerations. This report presents a landscape review of issues specific to RD patient registries for research. It summarizes current guidance for designing new RD registries for research to improve patient outcomes, giving special attention to guidance that differs from that for registries of common conditions and to guidance on maximizing the validity and representativeness of registry data.

Our search did not indicate that there are requirements for registering studies of RD. There are proposed regulations to require registration of clinical trials (for rare or common diseases) at clinicaltrials.gov.

### Methods

Few systematic evaluations and peer-reviewed publications or recommendations on practices for RD registries exist. Thus, the guidance described herein reflects current practices and opinions ascertained from the peer-reviewed and other literature. Because this was a landscape review, we did not systematically identify and evaluate all reports relevant to our objective. Instead, we searched PubMed using predetermined and ad hoc search terms and reviewed

references from the Advisory Working Group, which was composed of Advisory Panel on Rare Disease members and PCORnet investigators. We reviewed the abstracts of identified publications and selected articles that seemed most relevant and informative. We included relevant publications with which we were familiar and germane references from identified literature. Lastly, we carefully reviewed websites that were referenced in publications, that we identified by Internet searches, and those with which we were already familiar. Whenever we encountered a promising reference, regardless of the source that cited it, we reviewed it. We describe our methods for this landscape review in detail in Appendix A.

## Requirements of Registries for Rare Diseases

The first step in designing a registry should be to establish the purpose and expected uses of the registry, because the purpose will determine the registry's population and data collection requirements. Some common reasons to establish a registry are to determine the natural history of the disease, to conduct surveillance for adverse effects of treatments, and to ascertain and recruit individuals for research studies. Other purposes of registries and how they influence registry design are discussed below.

Once the registry developers have determined the purpose, they should review existing registries to confirm that no other registry or dataset can fulfill the purpose. Several organizations compile lists of registries that can be used to identify existing registries targeting a specific RD (Table 2). Establishing a new registry that draws from the same population and has the same purpose as an existing registry is inefficient and may compromise the representativeness of both the existing registry *and* the new registry (Workman 2013). If existing registries target the same RD as the proposed registry, but draw from different populations, comparing data between the new and existing registries may be of value. Shared methods facilitate such comparisons. In addition, insights from existing registries in successful and unsuccessful methods may be useful when designing new registries.

*Table 2. Directories of Disease Registries*

| Organization | Comments |
|---|---|
| Agency for Healthcare Research and Quality | In addition to listing existing patient registries, serves as an archive for expired registries |
| National Institutes of Health | Lists only national registries |
| Orphanet | European RD registries |
| RD-Connect | Global consortium of RDs. Includes a directory of member registries |

Although RDs have similarities, there is also considerable variation among them. Many RDs are the result of mutations in single genes, but others are partially or solely the result of environmental cases. The characteristics of the specific RD may affect many aspects of

organizing and using a RD registry, including recruitment strategies, sources and potential magnitude of bias, analytic strategies, and the generalizability of the registry findings to other populations. Some disease characteristics that may impact the design of a RD registry include:

- the etiology of the disease
- the incidence and prevalence of the disease
- the natural history of the disease and typical age of onset of symptoms
- How heterogeneous the disease is
- the outcomes of interest and how frequently do they occur

### *Purpose and Anticipated Data Uses*

RD patient registries can inform clinical and public health practice and serve as a source of research data and of cases for recruitment of more targeted research. Some specific uses of registries and registry data are discussed below.

### *Monitoring the Natural History of a Rare Disease*

A registry can provide essential data to describe the natural history of a RD. A critical consideration when designing a registry for this purpose is the representativeness of disease course of the registry enrollees. Three factors that can affect the representativeness of disease natural history of the registry enrollees are the completeness of case ascertainment, the timing of diagnosis in the study population, and the potential aggregation of cases near facilities that provide specialty treatment or services for individuals with the RD.

Ensuring that the registry includes every case of a RD within the study population may be very difficult. Completeness of case ascertainment is less of a concern if the cases in the registry are representative of all cases with the RD within the study population. In most cases, the representativeness of a registry must be inferred from the methods of case ascertainment and the attributes of the cases.

Ascertaining cases only from clinics that specialize in a particular RD will omit any cases who receive all of their treatment elsewhere—either because they lack the resources needed to obtain care at the specialty clinic or because they prefer to receive care elsewhere. If patients with a RD or their families move to be nearer to facilities with specialty services or treatment, the prevalence of the RD will be higher near the facilities, and conversely, lower in areas some distance from the facilities. The prevalence estimated based on registry data may be higher or lower than average if the concentration of such specialty facilities within the registry catchment area is unusually high or low. For example, the North Carolina MD STARnet surveillance area encompasses the central region of the state and includes all four muscular dystrophy clinics

within the state. If families of boys with Duchenne muscular dystrophy (DMD) move from the western or eastern parts of the state to be closer to these clinics, the prevalence of DMD within the surveillance area will be higher than would be expected based on the size of the population and the incidence of DMD.

For RDs without unique or distinctive symptoms, the path to diagnosis can be lengthy and unpredictable. Diagnosticians usually do not consider a RD diagnosis until ruling out more frequently occurring conditions or until a distinctive constellation of symptoms emerges. The lengthy path to diagnosis that occurs for many RDs—the "diagnostic odyssey"—results in cases being diagnosed at differing stages of their natural history (Hilbert et al. 2013; Wong et al. 2015). The path to diagnosis for a given disease may differ by patient characteristics if these characteristics are related to the ability to obtain care or in-depth diagnostic testing or to provider perceptions of who is at risk for a specific disease.

### *Improving Clinical Practice*

Registries can provide data on associations between specific clinical practices and patient outcomes. These data provide evidence on the effectiveness of a therapy, identify the attributes of patients for whom a therapy appears most effective, and identify specific clinics whose patient outcomes appear better or worse than those of other clinics (Quon and Goss 2011). An example of such a registry is the Cystic Fibrosis Foundation registry, which was established in 1966 and is credited with driving improvements in the health and survival of patients with cystic fibrosis in the United States over the past 50 years (Quon and Goss 2011).

### *Post-marketing Surveillance*

The U.S. Food and Drug Administration (FDA) has statutory authority to require postmarketing surveillance of new pharmaceuticals or devices, often referred to as Phase 4 testing (Crowther 2013). Postmarketing surveillance of patients who are receiving a new drug or other treatment is common and is especially important for therapies for RDs, which—on average—are tested on fewer individuals than therapies for common diseases (Bell and Tudur Smith 2014). Less common adverse effects of the treatment may not have been seen in the initial safety testing or clinical trials, which may have had small numbers of individuals with the RD. Also, the effectiveness of a treatment in clinical practice may be less than was observed in clinical trials (Wolfe and Michaud 2010). If postmarketing surveillance is a desired use of the registry, its designers may want to consult pharmaceutical companies or FDA officials in the registry's design.

### *Monitoring Patient Experiences*

Registries can collect data on patients' experiences of living with a RD. These data can inform the development of needed public health or social services or changes in policy to reduce barriers to treatment or improve independence. For example, registry data may allow

assessment of the length of time to diagnosis or the rate of misdiagnosis and the impact on receipt of effective therapy. Patients with a RD may be misdiagnosed and receive treatment for the wrong condition (Hilbert et al. 2013; Montserrat Moliner and Waligora 2013). Delayed diagnosis may postpone receipt of therapies that could slow disease progression, such as corticosteroids for children with DMD (Moxley et al. 2010). Registries may also provide data on factors that delay correct diagnosis, such as lack of genetic testing, or allow monitoring of the receipt of appropriate clinical care or adherence to clinical guidelines. If clinical care is inadequate or noncompliant, registry data or associated studies can be used to investigate public health policies or programs that could improve the care received, such as reimbursement for transport to a specialty center.

#### *Recruitment for Additional Research*

Registries can serve as a source for recruitment of survey and clinical trial participants, greatly increasing the efficiency of case ascertainment and recruitment (Malek et al. 2014). The usefulness of registries for this purpose depends on the representativeness of the registry participants and the ability of the registry's managers to provide information on its representativeness to researchers.

### Operations

#### *Case Definition*

The case definition for eligibility in the registry should reflect the purpose of the registry. The case definition for a RD patient registry should be based on a disease or group of diseases, rather than the receipt of a specific therapy or intervention (National Organization for Rare Disorders (NORD), Rare Diseases Europe (EURODIS), & Canadian Organization for Rare Disorders (CORD), 2012). Therapeutic- based registries are less comprehensive and may not reflect the full spectrum of the disease because of difficulty in obtaining or affording the treatment or limitations on receipt of the intervention.

If the disease is variable in its age of onset or symptoms, registry inclusion may be restricted by the age at diagnoses or the onset of disease symptoms to increase the disease homogeneity of the participants or to manage workload and costs. For example, a population-based registry of Duchenne and Becker muscular dystrophies (DBMD) excluded cases who did not manifest signs or symptoms by age 21 years (Miller et al. 2006). Registry designers may also choose a broad case definition to ensure that the full range of clinical manifestations associated with etiology are represented in the registry. Designers of registries for genetic RDs should consider whether the case definition should include requirements regarding genetic testing or the presence of specific genetic mutations. The Human Variome Project may provide insight on the range of disease mutations known or suspected to cause the targeted RDs (AlAama et al. 2011). Other

factors to consider as inclusion or exclusion criteria include the following (Richesson and Vehik 2010):

- **Diagnostic specificity.** Diagnostic methods change over time and the level of diagnostic investigation varies among cases. Registry designers may limit enrollment to individuals with a conclusive diagnosis, such as requiring genetic testing or using inclusive criteria. If more inclusive methods are used, the method(s) and date(s) of diagnosis should be collected. Researchers may choose to limit some studies to specific methods of diagnosis. For example, a DBMD registry categorized cases as "definite," "probable," "possible," "asymptomatic," "female," or "not a case" based on the availability of diagnostic test results (Mathews et al. 2010). Many researchers have limited their studies from this registry to cases with definite or probable DBMD diagnoses (Ciafaloni et al. 2009).

- **Clinical symptoms.** Registry designers may limit inclusion to cases who manifest specific clinical symptoms or a specified level of disease severity.

- **Geographic area.** Depending on the purpose(s) of the registries, designers may restrict the geographic scope to an area where they can reasonably achieve their ascertainment goals or to a well-defined population (National Organization for Rare Disorders (NORD), Rare Diseases Europe (EURODIS), and Canadian Organization for Rare Disorders (CORD) 2012).

- **Demographic characteristics**. Designers may limit inclusion into the registry to focus on the registry's primary purpose or for logistic reasons. For example, a registry of an X-linked genetic disorder focused on the disease course in affected males may exclude female patients, even if they are symptomatic. Or a registry with the goal of describing the natural history of a rare disorder in a minority group may restrict the registry to that minority group. If a major change in diagnostic methods occurred, such as the identification of a gene or development of a diagnostic test, designers may restrict inclusion to cases born before or after a specific date.

### *Case Ascertainment and Data Sources*

Recruitment or case ascertainment can be passive or active. Passive ascertainment solicits patients or families to enroll in the registry or requests clinicians to report cases. Solicitations may be made through service organizations, advocacy groups, websites, or mechanisms (Allen et al. 2008; Horton, Mehta, and Antao 2014; Johnson et al. 2014; Malek et al. 2014). One RD registry developed an Internet-based, open-source registry to recruit and enroll affected individuals (Bellgard et al. 2012). With active ascertainment, a registry seeks to identify all cases within a specified population through sources such as medical care facilities or administrative data.

Passive ascertainment requires fewer resources than active ascertainment, but the completeness and representativeness of the included cases cannot be determined. Recruiting

participants may require publicizing the registry and inviting individuals with the RD to participate. For example, the National Registry of Veterans with amyotrophic lateral sclerosis (ALS) recruited participants through publicity to associations including the ALS Association, American Academy of Neurology, and veterans associations (Allen et al. 2008). Social media has been effective in promoting participation: In 2012, an online registry for neurofibromatosis type 1 recruited 880 participants, 72% of whom became aware of the registry through Facebook (Johnson et al. 2014). The validity of patient-reported diagnosis is often considered a concern, but at least two reports have documented the validity of participant-reported diagnoses in Internet registries (Allen et al. 2008; Sharkey et al. 2014).

In addition to resource concerns, difficulty in gaining access to records, loss of records, and incorrect coding of diagnoses are major challenges for active ascertainment. Further, cases of interest are usually identified through health coding systems such as International Classification of Diseases, version 9 (ICD-9), ICD-10, or Systematized Nomenclature Of Medicine–Clinical Terms (SNOMED CT). Many RDs do not have specific codes, making computerized identification of cases difficult (Fung KW 2014). Nonetheless, researchers have successfully identified developed algorithms to identify patients with calciphylaxsis, a RD that lacks a unique code (Nigwekar et al. 2014) and to assess the accuracy of coding for ALS (Kaye, Sanchez, and Wu 2014). Their success suggests that algorithms could be developed to make or assess the accuracy of other RD diagnoses. We did not identify any reports of ascertaining cases by electronically scanning text fields of electronic medical records, the use of which became mandatory in 2015. Administrative data, such as hospital discharge summaries; insurance records, including Medicaid and Medicare; and birth and death certificates may be useful for case ascertainment (Kaye, Sanchez, and Wu 2014; Nigwekar et al. 2014). In many cases, identifying information and specific diagnoses cannot be obtained.

Registries that require recruitment or consent may be challenged in meeting their participation and retention goals because of distrust of researchers, especially among racial minority populations (Ford et al. 2005). The attitudes of health care providers toward the benefits of their patients' participation in a registry are likely to influence patients' willingness to participate (Ford et al. 2005).

### *Data Collection*

Data collection methods relate to, but do not completely overlap with, the case ascertainment method. When cases are ascertained passively, the source provides at least some of the data, with the nature and extent dependent on the type of source (e.g., patient, caregiver, or clinician). Patients may also be asked for consent to review their medical records. When cases are actively identified, data are collected from the ascertainment source. Additional data may

be sought from other sources, including the patient or family, primary care physician, or administrative records, although response from clinicians may be low.

For clinical data, medical records from clinical specialty clinics have the most complete data; vital records and insurance claims have the least complete data. Online, mailed, or telephone questionnaires can gather information from patients or caregivers that is not available in medical records or administrative datasets. Such information can include patient involvement in decision making and other patient-centered outcomes measures, and barriers to care, quality of life, and other nonclinical outcomes. For this reason, one of the 10 key principals for RD registries identified by RD advocacy groups was "Rare disease patient registries should include data directly reported by patients…" (National Organization for Rare Disorders (NORD), Rare Diseases Europe (EURODIS), and Canadian Organization for Rare Disorders (CORD) 2012)). For longitudinal registries, to the extent allowed by available resources, the frequency of follow-up data collection should reflect the registry's purpose, the rate of progression of the RD, and the introduction of new therapies.

### *Data Elements*

A registry's purpose dictates the domains of data to be collected (Gliklich, Dreyer, and Leavy 2014), but advocacy is increasing for including a minimum set of common data elements (CDEs) in all patient registries (National Organization for Rare Disorders (NORD), Rare Diseases Europe (EURODIS), and Canadian Organization for Rare Disorders (CORD) 2012). CDEs include standard variable definitions, code lists, and instructions that are applied across studies and registries so that the data are comparable (Gliklich, Dreyer, and Leavy 2014; Grinnon et al. 2012). CDEs may apply across disease or therapeutic areas or be disease specific. CDEs may reduce the effort needed to develop a database and enable registry data to be more easily linked or compared with data from other studies. They promote standardized data collection and improve data quality (NINDS Common Data Elements 2015). Using CDEs may lower the cost of developing a new registry, making registries more feasible when funding is limited, and may enable data from multiple small registry projects to be linked or compared to increase knowledge (Gliklich, Dreyer, and Leavy 2014). Several sets of CDEs for RDs exist, including the following:

- Core CDEs with Domain-specific (GRDR® CDEs) developed by NIH as part of the GRDR® Program (Rubinstein YR and McInnes P., 2015); and
- CDEs developed by the EPIRARE project for the European platform for RD patient registration (Taruscio et al. 2014); and
- The French national Minimum Data Set for Rare Diseases, which are very similar to the CDEs developed for the GRDR® (Choquet et al. 2015).

Several NIH Institutes have developed CDEs or have CDC initiatives ongoing. For example, the National Institute of Neurological Disorders and Stroke (NINDS) CDE Project developed core CDEs for studies of neurologic diseases in general, and supplementary CDEs for specific neurologic diseases (Grinnon et al. 2012). The NIH CDEs Working Group has developed a data base of all CDE initiatives within NIH, federal agencies, and other organizations (U.S. National Library of Medicine 2013). Although RD registry developers may find some of these CDEs useful, they were not developed specifically for registries or RDs. In contrast, the GRDR® and EPIRARE CDEs aim to be comprehensive for data needed for RD registries. We recommend that registry developers start with the set of variables in either GRDR® or EPIRARE CDEs and, if needed, supplement with standardized CDEs from other fields. These CDEs include validated patient-reported outcomes. Table 3 lists the GRDR® domains and example CDEs. Information about the specifications and anticipated uses of the CDEs in each domain are available online ("Global Rare Diseases Patient Registry Data Repository,") or in print (Taruscio et al. 2014). Note that in Table 3, R signifies that the CDE is required for the dataset, "O" signifies that the CDE is optional for the dataset; and N signifies that the datum is not part of the dataset.

*Table 3. Recommended Domains and Common Data Elements for Rare Disease Registries*

| Domain | Example CDE | Source | |
|---|---|---|---|
| | | GRDR® | EP |
| Current contact information | Person's last name[1] | R | R |
| | Did person consent to registry inclusion? | R | R |
| | Person's address | R | R |
| Sociodemographic | Person's date of birth | R | R |
| | Has the person died? | R | R |
| | Does the person have health insurance? | O | O |
| Diagnosis | What is the person's diagnosis? | R | R[2] |
| | Age when symptoms began? | R | R |
| | What test(s) to make diagnosis?[3] | R | R |
| Family history | Which blood relatives have same RD? | R | R |
| Birth and reproductive history | Person's birthweight | O | N |
| | Person's number of live births | O | O |
| Anthropometric | Person's recent weight? | R | R |
| | Age of recent weight | R | R |
| Behavioral health[4] | Current tobacco use? | N | N |
| | Frequency of having ≥6 drinks on one occasion? | N | N |

(continued)

*Table 3. Recommended Domains and Common Data Elements for Rare Disease Registries (continued)*

| Domain | Example CDE | Source | |
|---|---|---|---|
| | | GRDR® | EP |
| Patient-reported outcome | Frequency feeling tired? | R | R |
| | Does the person's health limit his or her vigorous activities? | R | R |
| Medications, devices, and health services | What medications is person currently taking? | R | R |
| | Does person use medical foods or have a special diet? | R | R |
| Clinical research participation and biospecimens | Has person been in ≥1 clinical trial? | R | R |
| | Has person donated a biospecimen? | R | R |
| | Where (hospital/clinic) was biospecimen donated? | R | R |
| Contact and communi-cation preferences | What is person's preferred way of contact? | R | R |
| Administrative | What is the person's GRDR® ID? | R | R |
| | Source registry | R | R |
| Outcomes | Occupational status | N | R |
| | Patient HRQoL Index Score | N | R |
| | Comorbidity | N | R |

Used by the registry to generate the Global Unique Identifier (GUID).

Diagnosis recorded in GRDR® and EPIRARE by selecting one of a prespecified list of RDs, which are mapped to a SNOWMED-CT code.

Each registry predefines a set of diagnostic testing responses that are presented in the survey. Also included is "None."

Data for this domain are not collected by GRDR® or EPIRARE, but are shown here as examples of CDEs developed by other NIH institutes.

NOTE: EP = EPIRARE indicators for the European platform for RD registration (Taruscio et al. 2014); GRDR® = Global Rare Disease Patient Registry ("Global Rare Diseases Patient Registry Data Repository,"); "R" signifies that the CDE is required for the dataset; "O" signifies that the CDE is optional for the dataset; blank = datum not part of dataset.

The registrar should review, update, and adopt new registry items as needed to reflect changes in the registry's purpose, sponsorship, or technological infrastructure. The registry developers should establish frequency of review as part of the governance of and protocol for the registry.

### *Data Quality*

Two kinds of errors can compromise the data that a registry collects. The first relates to the representativeness of people in the registry; the second relates to the validity, accuracy, and completeness of the data that the registry collects (Richesson and Vehik 2010). Furthermore, these errors can be random or systematic. Random errors are ubiquitous and nearly impossible to eliminate. The registrar's concern is minimizing their frequency. Of more concern is systematic error, which can introduce bias in study results. Data quality reports should be

generated frequently when a registry is first implemented, when problems are detected in the registry's conduct, or results or when the registry's methods change.

The following discussion describes some common types of random and systematic errors and methods for decreasing the likelihood of their occurrence.

### *Representativeness of Patients in the Registry*

Registry design or errors in case ascertainment can result in the registry having an unrepresentative sample of patients with the RD of interest. Self-selection bias occurs when affected persons who seek a diagnosis for their condition or agree to participate in a registry differ from those who do not. The bias can be countered by systematically searching all the sources in a population where an affected individual could be found, but even the most diligent active ascertainment is unlikely to identify undiagnosed cases. Lead-time bias results from systematic differences in the age or severity of morbidity of affected individuals at the time of diagnosis. This bias is a particular concern when disease onset is highly variable in timing and presentation, and when the likelihood of diagnosis is affected by patient characteristics, such as geographic residence. Such bias is difficult to surmount unless everyone in the study population is tested for the disease (e.g., newborn screening). Misclassification bias results from a tendency for patients with particular characteristics to be misdiagnosed.

### *Completeness and Accuracy of Data Collection*

Information errors occurs when a datum is not available, incorrectly recorded, or measured using devices that are not calibrated to the same standard. Information error can be random or systematic. In assessing whether missing data occur randomly or systematically, registrars need to carefully examine their data and, if possible, devise ways to collect or impute the missing data. Registrars can monitor data quality through automated data edits and frequent review of key variables.

### *Protection of Human Subjects*

In the United States, the conduct of a registry often must be approved by an Institutional Review Board (IRB) associated with a registrar or with the source from which the registrar collects data. In addition to considering the protection of research subjects, IRBs usually examine a registry's compliance with the Health Insurance Portability and Accountability Act (HIPAA). When a registry's data collection is deemed a public health activity (e.g., surveillance mandated by public health regulations), approval of the registry's protocol by the IRB may not be required (Centers for Disease Control and Prevention 2003). However, the rationale for this determination must be documented and submitted to the IRB.

Obtaining approval for the collection of de-identified data may be easier (Sengupta, Calman, and Hripcsak 2008). This option is not useful for longitudinal registries that must know an

individual's identity to collect or link data over time. Registries linked to biospecimen banks require informed consent (IC) of the participants. A prototype consent form has been developed collaboratively (Rubinstein et al. 2014).

Whether or not the registry requires formal IRB review, the designers should consider issues regarding the protection of individuals enrolled in the registry. These issues include whether the information returned to the participants is limited to aggregate data or whether their individual psychometric or laboratory test results or clinical evaluation findings will be returned to participants upon request. Registries that conduct or include genetic testing results should consider the recent recommendations from the American College of Medical Genetics on the return of incidentally found clinically relevant genetic findings (Green et al. 2013). Another issue to consider is whether participants will receive compensation for participation and the structure of any compensation or benefits to be provided.

#### *Modifying the Registry*

Once established, a registry may require modification if its purpose, sponsor, or the technical infrastructure that underlies its operation changes or if operational problems are detected (Gliklich, Dreyer, and Leavy 2014). Registry modifications require assessments similar to those needed for the original design and implementation.

### Registry Management

#### *Governance*

The governing structure of a registry is determined by its sponsor, its purpose, and its stakeholders. The stakeholders are the individuals and groups invested in the success of the registry and committed to its purpose(s), function(s), and success. Stakeholders typically include regulatory authorities, clinical care providers, public health practitioners, manufacturers of therapies, researchers, advocacy groups, patients, and their families (Aymé S 2011). The sponsor of a registry provides or obtains funding for the registry and may host or operate the registry. The sponsor and stakeholders determine the purpose and parameters of the registry. They may be represented by an advisory committee.

The advisory committee can serve many functions, but a common function is to represent the interests of the registry stakeholders. Its members may provide input on the purpose of the registry, its relevance to the stakeholders, and the engagement of stakeholders (Gliklich, Dreyer, and Leavy 2014). It may set registry policy on ethics or data access, use, and stewardship; or oversee the administration of the registry and monitor its financial, clinical, and social sustainability (Montserrat Moliner and Waligora 2013).

### *Administration*

The registry is administered by a registrar, who has primary responsibility for the design and conduct of the registry. The registrar and his or her staff create, maintain, and implement the registry's protocol; maintain the database; promote its analysis; and arrange for its evaluation. The registrar is responsible for stewardship of the registry's data, including implementation of its data access policies. Often, as analyses of registry data are reported in presentations at professional meetings or published in professional journals, understanding of the registry's data grows and demand for them increases.

### *Registry Software*

As more registries for RDs are developed, the need for interoperability among them becomes increasingly obvious (Forrest et al. 2011). Interoperability is facilitated substantially by the availability of free software for the infrastructure of a web-based registry for RDs, such as that available from NORD (National Institutes of Health 2012) or the software that was developed by the Marshfield Clinic Research Foundation and is distributed by the NIH/NCATS GRDR® Program to support data sharing with the GRDR® Program. Another software which is a second-generation RD registry framework permits customized data elements (Bellgard et al. 2013; Bellgard et al. 2014).

Registry designers can also design the registry such that it can be converted to a common data model (CDM), such as the Observational Medical Outcomes Partnership CDM (Overhage et al. 2012). CDMs allow the same analyses to be run against multiple datasets with minimal modification, greatly increasing the feasibility of combining or comparing data from different registries. A full discussion of CDMs is beyond the scope of this review; registry designers may wish to consult an informatics specialist to take advantage of recent developments in this area.

### *Data Access*

Registry staff and investigators are unlikely to have the capacity for conducting all analyses of interest on registry data. Data access portals that allow simple or complex data queries or that allow investigators to request data for analysis can greatly increase use and impact of the registry data. Examples of such portals are the Orphanet portal (Orphanet 2014), which provides information on RD research, orphan drugs, and other topics, and the GRDR® repository, which integrates data from all types of RD registries to be available for cross-disease analyses and various biomedical studies (Rubinstein et al. 2010).

## Part 2 Stewardship of Biospecimens and Biobanks

As discussed above, many registries collect biospecimens from their enrollees. A biospecimen is a quantity of tissue, blood, urine, or other human-derived biological material. It can comprise

subcellular structures, cells, tissue (e.g., bone, muscle, connective tissue, and skin), organs (e.g., liver, bladder, heart, and kidney), blood, gametes (sperm and ova), embryos, fetal tissue, and waste (urine, feces, sweat, hair and nail clippings, shed epithelial cells, and placenta). Portions or aliquots of a biospecimen are normally referred to as samples (National Cancer Institute 2011).

A collection of human biological specimens and associated data that is stored in an organized system is referred to as a biobank. The biological materials are not only annotated with medical information (health records, family history, images) but more than likely include epidemiological data (e.g., environmental exposures, lifestyle/occupational information). Specimens and associated data are usually coded or anonymized to ensure the privacy of the donor but may have the ability to link back to the donor to provide relevant clinical information. Biobanks can be typically found at international, national, and local levels and may vary in size, scale, scope, and type (O'Brien 2009; Parodi 2015).

Biospecimens are expensive to collect and maintain. Good stewardship of the biobank and its specimens is critical to maximize the value obtained from the specimens and protect participant privacy. Stewardship implies a more active role in the handling of biospecimens than the passive characterization of custodianship. To be a steward means not only being responsible and accountable for the preservation of the specimens and data from the time of collection through research use but also the ability to actively promote and foster the sharing of the biospecimens and associated data within the scientific community so others may derive research value. Overall, the foundation of stewardship is built on careful planning and policies that ensure long-term quality of the biospecimens, and the confidentiality of associated data, privacy of the participant, and the agreed use of the specimens as implied in the signed IC. Biobanks and their procedures, such as IC requirements, are monitored by IRB committees to protect the rights of the donor and stakeholder interests (Lowrance 2012; O'Brien 2009).

The main focus of a biobank is to collect, process, store, and distribute the highest quality biological materials for medical research and to make the specimens and associated data available to the widest possible range of scientific research. The quality of biospecimens is directly related to the validity and completeness of the associated biospecimen data profile. A biospecimen profile usually includes the following: patient demographics and medical history; biospecimen collection and processing details; storage procedures; the type, nature, and composition of the biospecimen; data yielded by analyses; and quality control data for both specimens and clinical data radiological, pathological-imaging, and clinical laboratory data. The detail of information is limited only by the available technology to capture, store, and integrate it and by the scope of the ethical and legal framework within which it is permitted to be used (Riegman et al. 2008).

The full optimal value of any biobank can be realized only in a climate of cooperation and sharing of resources. The ultimate goal of a biobank is to increase the quality of patient care and hasten the impact of research on the care itself. Biobanks can vary in scope, ranging from formal government, academic, and commercial organizations to informal collections of materials in an individual researcher's freezer. Biobanks are heterogeneous, although they do have some commonly shared operational characteristics. However, biobanks for the most part have distinguishing traits that can directly affect their scope; these include size, research design, types of samples collected and stored, collection methods, donor recruitment, informatics support, consent procedures, and governance structure. The following biobank designs support current medical research projects (Gottweis et al. 2012):

- **Population-based Biobanks.** These typically recruit healthy donors who are representative of a region, country, or specific ethnic group. The main goal is to discover biomarkers for disease susceptibility within a definite population.
- **Disease-oriented Biobanks.** Biological materials found in such biobanks are usually collected from patients within the context of clinical care. Patients can be resampled at follow-up visits during the course of their disease treatment.
- **Case-control Biobanks.** A prerequisite for meaningful case-control studies is the collection of matched (age and sex as a minimum) individuals presenting a given disease with compatible healthy controls.
- **Tissue Banks.** Tissue banks can represent diverse collections of tissue specimens or specimens of the same type of tissue (e.g., brain, lung) from either living or deceased patients. The tissues are usually cryopreserved or chemically preserved. Formalin-fixed paraffin embedded specimens are a common material type found in these collections.
- **Other Types of Biobanks.** Residual samples from clinical trials can be integrated into a biobank design and used in research. These types of samples are usually accompanied by valid and complete clinical and laboratory data profiles. Other types of specimens can include Guthrie cards from neonatal screening programs to detect congenital disorders and umbilical cord blood for use in therapeutic transplants.

### Sample Types

Biobanks can collect many types of samples depending on the purpose of the biobank. Individuals whose data or biological materials are used in research may agree to provide a biological sample for a particular project while they are living or donate organs, tissue, or their entire body for research after their death. Samples may be collected through routine clinical procedures or through additional medical procedures, when needed. Research involving human

biological materials may include any or all of the following: tissues, organs, blood, plasma, skin, serum, DNA, RNA, proteins, cells, hair, nail clippings, urine, saliva, or other bodily fluids (Maschke 2008). Human biological materials may be obtained by researchers for several reasons (Canadian Institutes of Health Research 2014):

- as a specific research purpose;
- medical or diagnostic procedures with no initial intent to be used in research; or
- research, medical, or diagnostic purposes with some expectation that the materials may be used for future research, although the precise purpose may not be known at the time of collection.

Biobanks are essentially a bioresource that are developed for current and future biomedical research purposes. They typically store a variety of human samples and associated data that are obtained from diseased and nondiseased populations. The specimens and data are sometimes stored for many years and eventually become available for use by third-party researchers through a well-defined application process. Investigators who design research studies that will use the storage services of a biobank should include the broadest possible language in their consent document that not only protects a study participant's confidentiality and privacy but also allows the optimal use of the samples and associated data by third-party researchers (DeRenzo and Moss 2006).

### *Virtual Biobanks (Specimen Locators)*

An opposite approach to centralized biobanking is the virtual biobank. A virtual biobank allows for the electronic integration of specimen and associated data through a common data registry that can be accessed worldwide regardless of where the specimens were collected or are currently stored. It provides authorized researchers with the ability to review collected data without requiring access to the physical sample. A virtual biobank can be located in one physical location (e.g., hospital or research institute) that implements a common storage environment or it can be a network of multiple biobanks that have reached an agreement to follow the best practices of biobanking and accept minimum standards for data sharing (De Souza and Greenspan 2013).

A good example of a RD virtual biobank is the Rare Diseases Human Biospecimens/ Biorepositories database (RD-HuB) that is overseen by the Office of Rare Diseases at NIH (http://biospecimens.ordr.info.nih.gov/). The database consists of seven modules; (1) Repository, (2) Disease, (3) Specimen Type, (4) Anatomic Source, (5) Processing Method, (6) Storage Method, and (7) Imaging. These allow a researcher to conduct a search to obtain a list of all the specimens in the database. RD-HuB also provides a number of links to best practices for specimen collection; models and templates for IC and guidelines for handling human subject

material for research and treatment; and links to related articles and protocols and other useful information. Two other specimen locator resources have also been recently launched: the International Resource Locator (www.IRLocator.isber.org) established by the International Society for Biological and Environmental Repositories (ISBER) that serves as a catalog for types of specimens, and RD-Connect (http://rd-connect.bibbox.org/web/guest/welcome), funded by the European Union.

## Considerations

### *Ethical and Legal Considerations*

The importance of addressing current and future research questions and concerns over privacy of genetic and health information should never be understated. Thorough, open, and honest communication is essential for biospecimen collection and the use of collected samples and associated data.

**Informed consent (IC)** is the process that potential research subjects use to make a reasonable and informed judgment about their involvement in research study, based on the risks or benefits to them as individuals. Investigators use an IC document to obtain consent and clearly communicate the intended purpose of the research and the collection and use of specimens. To enable clear communication, consent materials should be given to subjects prior to the research visit so there is sufficient time for review. Study subjects should be encouraged to ask questions to confirm their understanding and their purpose for participating and what will be learned from the studies. The signature of the person conducting the IC discussion confirms that an understandable communication and dialogue has taken place.

The consent document should identify all intended uses of the biospecimens and associated information. If the research is sponsored by a commercial organization or has possible commercial intentions, this should be clearly described in the IC document and communicated to the study participant. Plans for archiving the subject's DNA or creating immortalized cell lines (which could provide an inexhaustible source of DNA for future studies) should be clearly revealed, and any plans for distribution of the subject's genetic materials to secondary users should be presented, even if such parties are not yet defined (Beskow and Dean 2008; DeRenzo and Moss 2006; National Cancer Institute 2011).

**Privacy.** An important factor in biospecimen research is protecting the privacy of individuals who contribute biospecimens and maintaining the confidentiality of associated clinical data and information. Applying the highest possible ethical standards is necessary to ensure the support and participation of human research participants, physicians, researchers, and others in biospecimen resource activities. With recent advances in genomic and proteomic technology, the sequencing of the human genome, and the increasing reliance of biospecimen resources on

electronic and web-based databases for data tracking, it is even more crucial to address the risk of breaches in privacy. The unintended release or disclosure of sensitive information can place individuals at risk for discrimination and related groups at risk for stigmatization, although the frequency of these types of harms is unknown (Cambon-Thomsen, Rial-Sebbag, and Knoppers 2007; Eder, Gottweis, and Zatloukal 2012; Gottweis et al. 2012).

**Legal issues** include adhering to relevant federal, state, and local laws and regulations surrounding the collection, storage, dissemination, and use of biospecimens; developing appropriate guidelines for biospecimen access; ensuring that biospecimens are used in scientifically meritorious research; and establishing biospecimen resource governance.

HIPAA, also known as "The Privacy Rule," set new standards and regulations to protect patients from inappropriate disclosures of their protected health information (PHI) that may affect a patient's access to insurance, employability, and privacy. Biobanks are legally and ethically obligated to protect data that are considered PHI. The increased demand for human specimens in genome-wide association studies and data sharing has raised concerns of privacy, confidentiality, and human subject protection. An important issue that may affect biobanks is the concept of providing research results to participants in studies. Most biobanks in the United States are not certified by the Clinical Laboratory Improvement Amendments (CLIA) and cannot legally provide study participants/patients with test results (National Institutes of Health 2012).

**Recontact.** It may become necessary to recontact research subjects to recruit them for future studies, or to obtain additional information or clarification for an existing study. Because it is important to protect the confidentiality and privacy of research subjects, it is normally not appropriate to recontact research subjects unless they had previously agreed to be recontacted when they consented to participate in the existing study. If an investigator anticipates a need to recontact a subject then a recontacting provision should be included in the IC document. This provision will allow subjects to "opt in" or "opt out" of future contact for an existing study or to participate in future research studies (Otlowski, Nicol, and Stranger 2010).

**Returning results.** The ethical decision to share research results and conclusions directly with study participants requires weighing the value of the disclosure against a benefit-based obligation. During the consent process a discussion should be held with the study participant about the expectation of receiving test results, especially if the results involve genetic testing. The following factors should be considered in determining the appropriate level of disclosure (Jarvik et al. 2014; Smith and Aufox 2013):

- Will you have the ability to identify and recontact participants, especially if data have been pooled?
- Are the results medically actionable?

- Can the results be validated in a CLIA-certified clinical lab?
- Could the interpretation of results change as more data become available and tests are applied to wider populations?
- Will the test result change because of technological advancements, especially those concerning genetic testing?
- Does your team have the expertise and resources to communicate the uncertain and interim nature of research results, including concepts of relative and absolute risks?
- Can you provide access to a health care provider to review the results?
- Can you provide adequate follow-up support for potential psychological impacts?

### *Logistical Considerations*

Creating a biobank is advantageous for a number of reasons. Sample quality can be maximized by using a centralized processing, storage, and distribution infrastructure for multiple studies. Costs for future studies can be monitored closely and potentially decreased through the use of existing infrastructure and informational systems and the integration of automated technology. The biobank itself provides researchers with a scientific, reliable resource for developing new study design methods to collect, process, and preserve specimens.

A number of logistical challenges and biological factors (e.g., training, temperature, time, shipping, endogenous degrading properties [enzymes, cell death]) can have a significant negative impact on the overall quality and potential future use of collected samples. To maximize the biological information that can be obtained from collected samples, it is incumbent on researchers to first understand what methods will be used to test the hypothesis, what type of samples are needed, and how the samples need to be handled, processed, and stored following collection (Vaught et al. 2011).

This can be accomplished by involving study staff, laboratory scientists, and biobank personnel in preliminary discussions long before the first sample is collected. The discussions should focus on how to best integrate specimen collection and processing methods with laboratory and biobank procedures. All crucial steps from beginning to end need to be identified and described in a workflow diagram. Then the entire process needs to be evaluated by reverse engineering all workflow activities. This process will help identify key quality control checkpoints to ensure that all steps of the process are workable and maintain a high degree of specimen quality.

**Storage.** A number of studies have carefully evaluated the potential changes in specimen quality that are directly influenced by transport temperature and storage conditions and could have a negative impact on specimen integrity and biomarker/analyte stability. Specimens may be

transported to a central location (biobank) or stored locally. Storage could include a variety of temperature conditions based on sample type, intended use, preservation method, and length of projected storage. Prior to specimen collection, temperature requirements for specimen transport or storage should be thoroughly researched to determine the best possible temperature conditions for maintaining the specimen for its intended use. If samples are to be stored locally, an in-depth review of the local storage capabilities should be performed to ensure that all specimen storage requirements can be handled according to study requirements (Shabihkhani et al. 2014).

**Quality control.** Sample quality control requires written procedures that define specimen labeling, barcoding, container selection, sample annotation, laboratory processing, testing, and storage. Based on the complexity of the collection process it may be necessary to plan and carry out a pilot study to assess and evaluate the workflow procedures and analysis methods being used to ensure that all steps in the collection process have been identified and offer the best conditions to preserve specimen quality and biomarker stability (Holland et al. 2003).

**Location management**. The ability to accurately identify, track, and locate samples (sample type, volume, properties, location, consent status, etc.) is a requirement. Each sample should be uniquely identified and labeled with a barcoded identifier. The use of participant names on collection tubes or devices should be avoided at all costs and times. Samples should entered into an electronic sample and storage management system that is simple to enter, locate, report on, and manipulate samples, otherwise biobank staff may become frustrated and develop nonapproved workarounds that could compromise system efficiency, security, and sample identification. To be usable, the system must be web-based, able to be used by multiple users, accept barcode scanning, have the ability to adapt to study collection and processing schemes, store consent, and meet all federal security requirements.

**Duration.** Biological specimens undergo numerous transformations following collection. These changes can cause denaturation of proteins, a redistribution of elements, and nonquantitative recovery of certain analytes from biofluids or tissues. The duration of specimen storage is defined as the period of time that a sample can be used after collection and preservation without significantly affecting the composition of the analyte being tested.

To maintain a high degree of quality during storage, the biospecimens should be processed and preserved as quickly as possible following collection. Appropriate volumes, concentrations, or size for aliquots and samples should be determined in advance of storage to avoid any thawing and refreezing of biospecimens. All unnecessary thawing and refreezing of frozen biospecimens or frozen derivatives (e.g., DNA/RNA) should be avoided. When thawing/refreezing is necessary, a biospecimen resource should follow consistent and validated protocols to ensure continued stability of the analytes of interest (Gillio-Meina et al. 2013).

Methods should be established to minimize disruption of the stable storage environment during sample retrieval. When selecting biospecimen storage temperatures, consideration should be given to the biospecimen type, the specimen material of interest, and the anticipated length of storage.

## Governance

### *Data and Sample Ownership*

Policies regarding intellectual property rights vary. Institutions (government, academic, and commercial) in the United States normally assert ownership rights over biospecimens stored in their repositories. However, some researchers and individuals who provided biospecimens for research have unsuccessfully challenged this ownership claim in court. For example, in Washington University v. Catalona (2008), the U.S. Supreme Court declined to review a biospecimen ownership case. The question was whether individual donors who provide biospecimens for research "retain an ownership interest allowing [them] to direct or authorize the transfer of such materials to a third party." The court of appeals said, "The answer is no."

The National Cancer Institute's (NCI's) "Best Practices for Biospecimen Resources" says that researchers and institutions should share research data and tools generated through use of biospecimens in a timely manner, and that biorepositories have no inherent rights to future intellectual property, such as reach-through rights to inventions made by using repository samples (National Cancer Institute 2011).

### *Data and Sample Distribution for Research*

To best serve the needs of the scientific research community, biobanks should establish guidelines for sample distribution and clinical data sharing that is consistent with ethical principles, governing statutes and regulations, and IC language. Requests for specimens should provide a scientifically sound and appropriate research plan. The following specific issues (as outlined in NCI's "Best Practices for Biospecimen Resources) should be considered by the biospecimen resource:

- Timely, equitable, and appropriate access to human specimens without undue administrative burden.
- Scientific merit and institutional research qualifications, proven investigator experience with the proposed method, and a research plan appropriate to answer the study question.
- Community attitudes and ethical/legal considerations as primary factors.
- Fair, transparent, and clearly communicated access procedures.

- Appropriate allocation of biospecimens based on the nature of the scientific investigation (e.g., discovery, prevalence, initial validation, and hypothesis testing) and the need for annotation. The level of identifiability of the biospecimen and related transfer documents should be appropriate for the proposed research.
- A mechanism for addressing disputes over allocation decisions.
- An investigator agreement covering confidentiality, use, disposition, and security of biospecimens and associated data.
- The parties' written agreement in an Material Transfer Agreement (MTA) or other appropriate document that is consistent, as applicable, with the NIH Research Tools Policy and other applicable NIH sharing policies (National Institutes of Health 2015).

### *Linkage to Clinical Data in Registry*

Access to clinical data sources is an essential component for biobanks. In a number of countries biobanks link to national health or other health-related databases to obtain clinical information on their participants. However, in the United States, a fragmented health system presents challenges to obtaining health data. Biobanks linked to large health systems or networks may have more complete medical information than other tertiary care centers. Linking the two data sets of biospecimen and patient medical information collected through registries can be facilitated by the use of the Global Unique Identifier (GUID). For RDs, the NIH/NCATS Global Rare Diseases Patient Registry Data Repository/GRDR® (https://grdr.ncats.nih.gov/) program was designed to advance research for RDs. The ultimate goal of the program is to improve therapeutic development and quality of life for individuals suffering with a RD. To protect patient privacy, only coded data are collected and stored using a GUID that is assigned to each patient's data.

However, it is possible to conduct studies based on longitudinal data using electronic health records (EHR) data as proven by research conducted in the electronic MEdical Records and GEnomics (eMERGE) network. The eMERGE network is an NIH-funded consortium of biobanks that are linked to electronic medical records, which have developed methods and conducted early-stage research demonstrating the usefulness of biobanks in translational medicine research. The eMERGE network currently comprises nine biobanks, including both adult and pediatric participants. eMERGE has developed tools for genomic research using EHR to select phenotypes and then share the phenotypes across the network (McCarty et al. 2011).

### *Support and Maintenance*

For more than a century, the collection and use of biospecimens have played a prominent role in research efforts to detect and study disease. Biorepositories or biobanks as they are now known provide a key focal point for the gathering and storage of biospecimens and their

associated data. Although the importance of biobanks is widely recognized, the development and creation of these bio resources face many challenges but none as important as financial sustainability. Historically, it is well known within the research community that not all biobank operations are successful. Some fail and disappear while others restructure through bankruptcy proceedings or are acquired by other organizations through mergers and acquisitions. The reasons for failure vary but mostly center on flawed marketing strategies, sustainable customer value propositions, or viable funding.

If an organization needs a biobank it should consider whether it is better to build one or outsource it. Several key areas of focus that organizations should take into consideration when first deciding the pros and cons of developing a biobank are listed below in Table 4. The table has been adapted from a table in Watson, 2014) and a PCORNET Guidance Document developed by The Biorepositories Task Force. Regardless of the decision to build or outsource, organizations must also remember that biorepositories are subject to regulations and are encouraged to follow industry best practices. The International Society for Biological and Environmental Repositories (ISBER) has developed "Best Practices for Repositories," which is currently in its Third Edition and reflects the collective experience of its membership (http://www.isber.org).

*Table 4. Access Level of Biobanking Readiness Questions*

| Areas of Focus | Questions to Assess Level of Biobanking Readiness |
|---|---|
| Mission, vision & strategic objective (Collis and Rukstad 2008) | • Mission—What is the underlying motivation for biobanking?<br>• Vision—What does the biobank strive to be in the future?<br>• Strategic Objective—Has a single goal for biobanking been developed that is measurable and time bound? |
| Availability of resources | • What current resources exist and what resources are needed to plan and manage a successful biobank?<br>• What unique scientific expertise is available?<br>• Is a biospecimen science resource accessible?<br>• Is business expertise available to develop a viable business plan?<br>• What economies of scale are present to provide value?<br>• Are research subjects and biospecimens readily available?<br>• Is there a community of researchers to use collected biospecimens?<br>• What technologies are needed and available for successful biobanking? |
| Organizational/ stakeholder requirements & structure | • Will biobank collection(s) support a single user, several research studies within one institution, or multiple users from multiple organizations?<br>• Does the infrastructure exist within the organization to support biobanking?<br>• What are stakeholder goals and motivations?<br>• What business model is most appropriate given this information? (centralized vs. decentralized; in-house vs. external vendor, or using an existing biobank)<br>• If in-house, who will manage and operate the biobank? |

(continued)

*Table 4. Access Level of Biobanking Readiness Questions (continued)*

| Areas of Focus | Questions to Assess Level of Biobanking Readiness |
|---|---|
| | • Who will develop and maintain resources (e.g., clinical databases, LIMS)?<br>• What teams will be involved (e.g., informatics, programming, faculty, PM.)?<br>• What biospecimens need to be collected? (Type of specimen, and disease focus)<br>• What value does the biobank have to the institution?<br>• What funding sources have been identified, and does the organization/stakeholders expect to contribute financially to the biobank?<br>• Is a fee-for-service model appropriate to fund biobank operations?<br>• Is there an expectation that the biobank will be financially self-sustaining? |
| Value proposition | • Are identified value metrics relevant to stakeholders and users?<br>• What is the societal value of the biobank?<br>• What scientific/research value may be derived from the biobank? |
| Efficiencies (Internal & external) | • What are the existing efficiencies that may be practically operationalized?<br>• Is it reasonable that costs for providing services be recouped if needed?<br>• What annotation of biospecimens is necessary, and can these data be efficiently obtained?<br>• Can users access the biobank and receive samples/data in a reasonable time? |
| Acceptability | • Who are the biobank' s stakeholders?<br>• Do public or private stakeholders trust in biobanks?<br>• Who will provide oversight (e.g., advisory board, community, ethics, legal)?<br>• Is the governance policy fair, allowing for transparent distribution of biospecimens?<br>• Is the biobanking of specimens viewed as acceptable by potential sample donors?<br>• Has the organization established communication about biobanking and received public input and representation?<br>• Do sample-sharing models meet NIH GDS policy, and organizational values? |
| Standards | • Is the biobank committed to sound and responsible best practices and standards?<br>• Have stakeholders demonstrated commitment to accepted standards of practice and quality approaches?<br>• If so, was this commitment communicated to user and donor groups? |

Fundamental business principles must be applied to the development and operation of biobanks to ensure scientific impact and long-term sustainability. The true costs of developing and maintaining operations must be clearly defined and include the market need for the particular type of biobank under consideration and understanding and efficiently managing the biobank's "value chain," which includes costs for case collection, specimen processing, storage management, sample distribution, and infrastructure administration (Vaught et al. 2011). The following list of business strategies found in the PCORNET Guidance Document can be adapted to either a startup or existing biobank.

- Develop a Strategic Business Plan - A solid and comprehensive plan should be written and revised annually, and may include these components:
  - Vision, Mission, and Goals – Include societal value

- Opportunity – including stakeholder needs
- Value Proposition (and value-added service offerings)
- Definition of Services – Define service offerings
- Competitive Analysis
- Business Development Strategy (Marketing Plan)
- Communication (Outreach) Plan
- Organizational Structure
- Management Team & Resource Identification
- Quality Assurance Procedures
- Capital and Resource Requirements (Operations Budget)
- Revenue Projections – May include grant, donor & service revenue if applicable
- Biospecimen Collection Targets
- Formal Continuity Plan – addressing possible operational disruptions
- Quality Assurance Procedures
- Performance Metrics – Desired measures of success - including societal value and research impact

- Develop an Implementation Plan including timeline, milestones, contingencies, and path to secure short-term funding. The identification of additional resources including teams and key players involved in day to day operations is required.

- Develop a Cost Recovery Model; a crucial means to maintain economic viability, and ensure both short and long-term financial support. Financial support is typically derived from a variety of methods, including public and private funding, grants, philanthropic donations, and contracts. In increasing numbers, biobanks are developing fee-for-service models, recouping operating costs by charging fees to researchers and industry a fee to access and utilize a biobank's biospecimens. Biobanks may also recoup costs by providing researchers with biobanking services for their collected biospecimens. The following information is typically utilized to develop a biospecimen fee schedule:

  - Revenue projections - Accurate revenue estimates from all anticipated revenue sources
  - Cost analysis – Total infrastructure and biobanking operations expenses (including collection costs) identified during each stage of biospecimen management. Specific costs should be identified for each biospecimen type and volume (if applicable).
  - Projected or historical biospecimen service data (i.e., the number of samples allocated per year)
  - Market data to determine typical industry charges for comparable biospecimens (if applicable)

# Part 3 Issues About Study Design and Strength of Evidence for Rare Diseases

## Introduction

Previous sections of the report focused on best practices for biospecimens and registries. Many research questions for RDs will require de novo identification of patients and prospective data collection. This section of the report describes potential study designs and strength of evidence approaches for RDs.

Research findings can represent a true relationship, a chance association, or a systematic error (bias) (Behera et al. 2007; Higgins, Altman, and Sterne 2011; Viswanathan et al. 2008). The best study designs minimize the effect of chance and bias. Some traditional strategies for these effects include the following:

- recruiting a sufficient sample to answer the study questions to reduce the effect of chance;
- randomizing and concealing allocation to reduce the risk of selection bias and confounding;
- concealing allocation, masking of participants and physicians, ensuring fidelity to protocol, and measuring and controlling unintended co-interventions within and across comparison groups to minimize the risk of performance bias;
- using intention-to-treat principles during analysis to minimize the risk of attrition bias;
- blinding outcome assessors and using valid and reliable outcome measures, applied consistently across comparison groups to minimize the risk of detection bias; and
- registering protocols ahead of analysis to reduce the risk of reporting and publication bias.

Studies of RDs have very specific constraints that limit the use of traditional strategies and designs to minimize the effect of chance and bias. Specifically, these constraints include difficulty recruiting an adequate sample size that is representative of the population, difficulty obtaining outcome data, heterogeneity of populations, and concerns about ethics and privacy. We first describe these issues in greater detail. We then identify specific study designs, their applications or adaptations to RD, and constraints or disadvantages. Finally, we describe strength of evidence systems.

## Special Study Design and Conduct Issues for Rare Diseases

### *Adequate and Representative Sample*

The most fundamental challenge to conducting an adequately powered study—thereby minimizing the effect of chance on the results of the study—lies in the small numbers of

patients experiencing the condition (Gagne et al. 2014). Another issue that compounds the problem of small numbers of potential recruits is their geographic dispersion (Gagne et al. 2014). An unrepresentative sample cannot support claims of inference to the larger population.

### *Availability of Outcome Data*

Measuring final health outcomes that occur only rarely can be a challenge even in non-RDs. Even if studies of RDs are successful in recruiting patients, they may still be underpowered, depending on the frequency of occurrence of the final health outcome (Gagne et al. 2014). When researchers seek information about rarely occurring outcomes in a RD population, they may need to rely on indirect chains of evidence (e.g., surrogate markers or evidence of the mechanism of action (Goodman and Gerson 2013) rather than direct evidence linking the intervention and the health outcome. In this instance, the risk of detection bias needs to be balanced with risk of random error.

### *Heterogeneity of Etiology, Presentation, and Course of Illness*

A single RD may be defined by its phenotype (outward manifestation), but it may include patients with a variety of genotypes (internal inheritable information) that interact differently with environmental factors (Venance, Herr, and Griggs 2007) and therapies. Within RDs, course of illness may vary from invariably fatal to relapsing-remitting. When these underlying sources of heterogeneity are known and measured, studies of RDs may find that issues of low power (and the risk of random error) are compounded by the need to account for heterogeneity. An even more challenging and common scenario is that sources of heterogeneity may be unmeasured, leading to the potential for confounding. One such concern relates to differences in access to care, which can influence the stage at which a patient with a RD may be given a diagnosis and then become eligible for treatment.

### *Concerns about Privacy and Ethics*

A significant concern for RD registries is the risk of loss of privacy (Mascalzoni et al. 2014). Interlinked registries and biobanks that share data to maximize the potential for research also risk re- identification of de-identified patients. The negative consequences of loss of privacy may extend beyond the patient to family members or even communities defined by race or ethnicity.

When registry developers and users interpret privacy as having the right to consent to access by third parties to the patient's own data, two assumptions come into play. One is that new studies using existing data must reconsent patients; the other is that patients have the right to withdraw at any time from studies. The current use of many registries is not consistent with these rights; strict application of the traditional consent process can significantly restrict the use of registries (Mascalzoni et al. 2014). Reconsent and withdrawal from registries have the potential to increase the risks of selection bias and confounding.

Relatedly, the Institute of Medicine has recently issued recommendations that support greater sharing of clinical trial data and has acknowledged that sharing must be balanced with respect for participants' privacy and right to consent (Institute of Medicine 2014). When privacy could be particularly vulnerable, as in the case of treatment for a RD, data sharing may need to include special safeguards such as decision making by an independent panel (Lo 2015).

Another concern that applies to both rare and non-RDs relates to the receipt of placebo, less effective, or ineffective treatments. Patients and physicians may seek to maximize the chance of receiving the more effective therapy based on prognostic factors; in observational studies and poorly randomized trials, these efforts can increase the risk of confounding. Similarly, physicians and patients may prefer existing and proven standard care over an untried experimental therapy (Day 2011). For diseases with few or no promising approaches other than the experimental therapy under study, patients may prefer to be offered any alternative that appears promising, rather than participate in studies with a placebo or control arm. Clinicians, likewise, may consider offering anything other than a potentially active treatment to be unethical. In the RD context, these considerations can serve to limit participation in trials that may be crucial to establishing efficacy.

### *Types of Clinical Research Study Designs, Applications, and Constraints*

Numerous publications lay out study designs for RD and describe their advantages and disadvantages (Bogaerts et al. 2015; Cornu et al. 2013; Gagne et al. 2014; Gerss and Kopcke 2010; Gupta et al. 2011; Institute of Medicine 2001; Korn, McShane, and Freidlin 2013; Tudur Smith, Williamson, and Beresford 2014).

Two publications provide an algorithm for choosing study designs (Cornu et al. 2013; Gupta et al. 2011) systematically reviewed the literature for RD research frameworks and study designs. They then generated an algorithm for choosing among one of six designs to address the issue of having a limited number of participants: crossover design, n-of-1 trials, response-adaptive randomization design, ranking and selection design, internal pilot design, and sequential design (Gupta et al. 2011). The algorithm poses questions relating to the predictability and duration of effect, stability of the disease course over at least two intervention periods, retention over at least two intervention periods, availability of the required number of participants, time between inclusion and outcome assessment compared with accrual time, and whether a planned sample size can be reasonably recruited (Gupta et al. 2011).

Cornu and colleagues, in a more recent publication (2013), looked at a larger list of possible designs that overlap in part with those offered by (Gupta et al. 2011). Specifically, Cornu et al. discussed parallel group, factorial, crossover, Latin square design, n-of-1, delayed start, randomized placebo-phase, stepped wedge, randomized withdrawal, early escape, three-stage, and adaptive randomization. The Cornu algorithm does not focus on sample size issues alone; it

considers issues such as reversibility of outcome, rapidity of response, whether time on placebo is minimized, whether active treatment is provided at the end of the trial, and whether controls are within (as with crossover designs) or across (as with independent samples of comparisons) patients (Cornu et al. 2013). Both algorithms focus primarily on randomized trial designs and acknowledge the use of "meta-methods" such as Bayesian analyses that could be used in combination with specific designs.

Table 5 describes each design, its potential application to RDs, and constraints. Although each design is listed separately, they are by no means mutually exclusive. Studies may combine multiple strategies to minimize the effect of chance and bias. For instance, all strategies listed under parallel-group design may be applied to other designs as well. Additionally, limited consensus exists on how to classify design types (van der Lee et al. 2010), as evidenced by the differences in the two available algorithms (Cornu et al. 2013; Gupta et al. 2011). Table 5 below attempts to be as inclusive as possible of proposed designs. Table 5 describes strategies for randomized designs, controlled clinical trials, and observational studies. No consensus exists on the risk of bias or hierarchy of evidence within each of these categories; as a result, the strategies listed under each section do not appear in any particular order. We also note an underlying requirement for all designs listed below: the sample must be a random selection of the population to draw inferences to the population.

*Table 5. Study Designs for Rare Diseases*[a]

| Design | Description and Application or Adaptation to Rare Diseases | Constraints |
|---|---|---|
| **Randomized designs** | | |
| Parallel-group randomized trial | • Patients stay with randomly assigned arm for duration of the study<br>• Lengthen trial duration to capture more events and thereby reduce sample size requirements among the trial participants<br>• Reduce heterogeneity in included patients by focusing on high-risk patients or by using genetic testing to select patients at high risk<br>• Create a factorial design in which multiple treatment comparisons are carried out at once, thereby reducing the sample size requirements for all questions (e.g., by evaluating the effect of combinations of interventions [A+B vs. placebo])<br>• When information on clinical endpoints is unavailable or rarely available, use continuous outcomes measure instead of a binary outcome, a surrogate marker instead of a hard clinical endpoint, a composite endpoint instead of multiple outcomes, or repeated measures instead of single measures; all these strategies can reduce sample size requirements | • Need discrete, multiple events per participant<br>• Need reliable and valid risk assessment tools or genetic tests<br>• Multiple treatment approach must be relevant and meaningful for RD<br>• No interaction should exist between treatments<br>• Sample size requirements are unchanged for questions about individual treatments (e.g., A vs. placebo, B vs. placebo)<br>• Alternate measures must be clinically meaningful (i.e., percentage reduction in continuous measures must represent a minimally important difference; surrogate markers must be closely and directly linked with health outcomes; composite endpoints must be valid and reliable; repeated measures should not be used if outcomes within individuals are likely to be correlated or if outcomes may be influenced by familiarity with the instrument [practice bias])<br>• Risk of loss of privacy with data sharing<br>• |

(continued)

*Table 5. Study Designs for Rare Diseases[a] (continued)*

| Design | Description and Application or Adaptation to Rare Diseases | Constraints |
|---|---|---|
| | • Use clinical trial networks and registries for RDs to help with recruitment of larger and more geographically diverse patient populations<br>• Integrate trials into clinical practice to enhance participation in studies; every patient is assigned randomly to a study arm | • Application of a traditional consent process requires reconsent, beyond the initial consent required to be part of a registry; this poses the potential for participants to choose to withdraw from the study, which can reduce the available sample and potentially create a selection bias<br>• Requires relaxing eligibility criteria, which may increase measured and unmeasured heterogeneity<br>• Requires equipoise or uncertainty of effectiveness among available interventions |
| Crossover randomized design | • Patients receive two treatments in sequence randomly, with a washout period between treatments<br>• Each participant serves as his or her own control, thereby reducing sample size requirements<br>• Latin square allows for multiple treatments in randomized sequence<br>• Each treatment appears only once in each sequence and treatment period | • Potential for detection and performance bias if effects of intervention carry over<br>• Requires short latency period for measuring a clinically relevant outcome<br>• Requires full effect of the intervention shortly after initiation and full loss of effect on termination of treatment<br>• Not applicable if effects of the intervention are irreversible, disease course is unstable, or effects of outcome influenced by order of interventions received<br>• Attrition may significantly undermine results when patients drop out before undergoing the crossover<br>• Potential for performance bias from period effect (effects attributable to the calendar time, e.g., season, in which the intervention was delivered)<br>• Same as for traditional crossover designs |

(continued)

*Table 5. Study Designs for Rare Diseases[a] (continued)*

| Design | Description and Application or Adaptation to Rare Diseases | Constraints |
|---|---|---|
| N-of-1 trial | • Single participant randomized to treatment(s) and placebo, in random sequence<br>• Each individual acts as his or her own control | • Potential for detection and performance bias if effects of intervention carry over<br>• Requires short latency period for measuring a clinically relevant outcome<br>• Requires full effect of the intervention shortly after initiation and full loss of effect on termination of treatment<br>• Not applicable if effects of the intervention are irreversible, disease course is unstable, effects of outcome influenced by order of interventions received<br>• Potential for performance bias from period effect (effects attributable to the calendar time in which the intervention was delivered) |
| Sequential randomized controlled trials | • Null hypothesis is tested in a series of interim or continuous analyses; these analyses then determine whether the trial should be terminated because of safety, futility, efficacy<br>• Variations include group sequential design (interim analyses at predetermined points) or boundaries design (continuous or group analysis mapped against *a priori* boundaries representing the balance between information gathered over the course of the trial and effect size, to determine whether the "sample path" stays within the boundaries)<br>• Trials allowing early termination require fewer patients | • Risk of incorrect rejection of null hypothesis (Type I error) because of multipletesting<br>• Potential for selection bias and confounding if participants and providers are aware of upcoming changes in design<br>• Requires short latency period for measuring a clinically relevant outcome |
| Adaptive randomized controlled trials | • Adaptive treatment allocation designs test the null hypothesis in a series of interim analyses; these analyses then influence subsequent randomization in the next phase<br>• Bayesian analyses (allowing updates of prior probabilities) or frequentist approaches can be used<br>• | • Disproportionate recruitment could reduce power<br>• Requires short latency period between intervention and outcome (which may be an activity biomarker or intermediate efficacy endpoint) for results to influence randomization in study<br>• Requires a binary outcome for defining success or failure<br>• May be suited more for exploratory analysis than for confirmatory analysis<br>• Potential for selection bias and confounding if participants and providers are aware of upcoming changes in design |

(continued)

*Table 5. Study Designs for Rare Diseases[a] (continued)*

| Design | Description and Application or Adaptation to Rare Diseases | Constraints |
|---|---|---|
| | • Adaptive treatment allocation designs allow the probability of being randomized to an intervention to change during the enrollment period; the probability of being randomized will increasingly favor the arm with the more promising results (play the winner) or increasingly penalize the arm with less promising results (drop the loser)<br>• It can also increase the proportion of patients assigned to the more favorable treatment, thereby increasing the number of willing participants | |
| Adaptive randomized controlled trials (continued) | • Adaptive designs can be used to narrow from a selection of doses (ranking and selection designs) rather than rejecting a null hypothesis<br>• Adaptive designs can be used to select among subpopulations and thereby balance covariates (covariate-adaptive randomization) and help address underlying heterogeneity | |
| Internal pilot | • Internal pilots allow data from participation on pilot trials to contribute to final results, unlike conventional pilots that can deplete available participants for a full trial because their participation serves an exclusion criterion<br>• Internal pilots reduce the required sample size | • Internal pilots offer little benefit if protocols require major change between the pilot and the full trial |
| Randomized placebo-phase | • Patients are randomized to varying lengths of exposure to placebo, but all patients receive treatment in the end<br>• Can be used for conditions with a rapid unfavorable evolution | • Power depends on the number of placebo variants |

(continued)

*Table 5. Study Designs for Rare Diseases[a] (continued)*

| Design | Description and Application or Adaptation to Rare Diseases | Constraints |
|---|---|---|
| Stepped wedge | • When interventions cannot be delivered to all patients at once, all participants start with control and then are randomly assigned over consecutive blocks of time to the intervention until all patients are on treatment | • Potential for performance bias because of contamination (i.e., providers apply treatment arm behaviors and services to control arm) |
| "Early escape" in randomized designs | • Patients can withdraw if they meet *a priori* criteria (per protocol) or by patient choice<br>• Early escape can be applied to various trial designs, including crossover and N-of-1 (Huang et al., 2014)<br>• May enhance study retention and power and reduce exposure to less favorable treatments | • High volume of early withdrawal could reduce power<br>• Requires a binary outcome for defining success or failure<br>• Requires focus on short-term outcome that occurs during intervention |
| Randomized withdrawal | • All patients receive active treatment, responders are then assigned randomly to placebo or treatment<br>• Minimizes time on placebo (only responders are allocated to placebo) | • Requires short latency period<br>• Potential for detection and performance bias if effects of intervention carry over into placebo phase<br>• Not applicable if disease course is unstable or has slow evolution |
| Three-stage trial | • Combines early escape (of nonresponders) and randomized withdrawal (of assignment of responders to subsequent placebo or treatment), so allows an opportunity to benefit from therapy, avoids treating patients who respond to placebo, and reduces exposure to unfavorable treatments<br>• Stage 1: Initial randomization to treatment or placebo<br>• Stage 2: Responders to treatment in stage 1 to placebo or treatment iStage 3: Nonresponders to placebo in stage 1 placed on active treatment; nonresponders in Stage 3 exit study and responders then randomly assigned to treatment or placebo | • Constraints of early escape and randomized withdrawal<br>• Additionally, risk of performance bias if washout period is not sufficiently long<br>• Risk of selection bias if participants barely miss the cutoff for responders and therefore miss active treatment<br>• Inappropriate if withdrawal of drug causes flare of disease greater than at baseline<br>• Not suitable for controlled assessment of safety |

(continued)

*Table 5. Study Designs for Rare Diseases[a] (continued)*

| Design | Description and Application or Adaptation to Rare Diseases | Constraints |
|---|---|---|
| **Controlled clinical trials (nonrandomized)** | | |
| Risk-based allocation (controlled clinical trial with randomized component) | • Low-risk patients are randomized to high-dose and standard treatment, high-risk patients receive high-dose treatment, thereby addressing concerns about the ethics of withholding treatment from high- risk patients<br>• A combined analysis allows the prediction of the added benefit of high-dose treatment | • Requires a valid and reliable delineation of high vs. low risk<br>• Requires that the intervention has a plausible dose-response effect<br>• Risk of performance bias from lack of masking in the controlled trial group |
| Delayed start (controlled clinical trial with randomized component) | • Patients randomized to intervention and placebo; after active control phase, all patients receive treatment, thereby addressing concerns about the ethics of withholding or delaying treatment<br>• Primarily useful for evaluating the effect of the treatment on symptoms or disease progression | • Risk of performance bias from lack of masking in the controlled trial phase<br>• Risk of detection bias from carryover effect or if treatment follow-up is not long enough to observe effect |
| **Observational designs** | | |
| Prospective inception cohort | • Inception cohorts limit participation to "new users," thereby avoiding selection bias and confounding as is common with prevalent users, whose response to treatment may be a function of prior therapies, course of illness, and so on | • Difficult to implement for RDs; identifying patients with RDs at inception may be difficult because of the potential time lag in accurately some diagnosing rare conditions |
| Case-control studies | • Case-control designs select known cases and matching controls from a larger cohort, thereby reducing sample size requirements, particularly for some rare outcomes of RDs | • Risk of selection bias and confounding if matching is not done appropriately |
| Cohorts with historic controls | • Comparison of prospectively treated patients with historic controls reduces recruitment burden for control arm | • Risk of selection bias |
| Pre-post design | • Patients receive usual care or standard intervention followed by tested treatment<br>• Requires a detailed understanding of the natural history of the disease to avoid issues of regression to the mean | • Potential for regression to the mean (natural improvement over the course of time is misattributed to the intervention)<br>• Risk of selection bias |

[a] Bogaerts et al., 2015; Cornu et al., 2013; Gagne et al., 2014; Gerss & Kopcke, 2010; Gupta et al., 2011; Hampson, Whitehead, Eleftheriou, & Brogan, 2014; Honkanen et al., 2001; Huang et al., 2014; Institute of Medicine; Korn et al., 2013; Lagakos, 2003; Tudur Smith et al., 2014; van der Lee, Wesseling, Tanck, & Offringa, 2008; van der Lee et al., 2010; Wang, Hung, & O'Neill, 2012.

In addition to the designs listed above, various analysis strategies can be employed for RDs, such as propensity scoring and instrumental variables. Propensity scores may be particularly useful when events are rare relative to the number of potential confounders. Bayesian analysis, meta-analysis, and decision modeling also offer ways of expanding on knowledge from a single trial through inference, pooling, and modeling.

## Types of Strength of Evidence Systems, Applications, and Constraints

### *Background*

In health care, systems related to grading the strength or quality of evidence included in a systematic evidence review refer to a transparent and structured process for presenting reviewers' confidence in their conclusions about the effects of drugs, procedures, and therapeutic interventions; the aim is to permit patients, clinicians, and policy makers to be able to use the results of systematic reviews effectively (Atkins et al. 2005; Guyatt et al. 2011). The Grading of Recommendations Assessment, Development, and Evaluation (GRADE) working group developed the most widely used system (Group 2014). The Agency for Healthcare Research and Quality (AHRQ) Evidence-based Practice Center (EPC) strength of evidence (SOE) grading approach was based on and incorporates GRADE methodology, emphasizing concerns commonly encountered by the EPCs (Berkman et al. 2014). The GRADE approach goes one step further than EPC guidance, providing direction on developing recommendations from systematic review findings (Andrews et al. 2013). Both GRADE and the EPCs have also developed separate guidance for evaluating diagnostic test performance (Schunemann et al. 2008; Singh et al. 2012).

No matter the grading system chosen or the types of study designs determined to be appropriate for inclusion as evidence, earlier steps in the systematic review process remain the same; these tasks specifically include critical strategies for limiting bias and ensuring the quality of the systematic review (Behera et al. 2007). These methods entail, first, defining the questions of the review and then specifying the patient populations, interventions, comparisons, outcomes, and settings that will be the focus. Based on these parameters, a systematic review involves conducting a thorough search of the available literature (Relevo and Balshem 2011) and then assessing the risk of bias for each included study (Cochrane Collaboration 2011; Higgins et al. 2011). One final step of any review, in preparation for assessing the strength of evidence, requires reviewers to synthesize studies that form the body of evidence to answer each key question either quantitatively using meta-analysis or qualitatively. A systematic review concerning the effectiveness of treatment for a RD would follow this basic structure, but it may have only a limited evidence base, comprising small studies that are less likely to be randomized controlled trials (RCTs) than reviews of more common disorders or therapies.

### *Considerations in Strength of Evidence Grading of Rare Disease Treatment Intervention*

In an SOE grading systems such as GRADE and EPC SOE, reviewers' confidence in their findings are based primarily on consideration of scores in five domains. These include study limitations, directness, consistency, precision, and reporting bias. Three additional domains that may be relevant and are sometimes considered are dose-response association, plausible confounding that would decrease the observed effect, and strength of association (magnitude of effect). Whether a domain is a particular concern in a review of treatment for a RD will depend on the study questions and characteristics of the studies included in the specific body of evidence. We separately discuss each domain below.

#### *Study Limitations*

Study limitations is a summary measure of the risk of bias of the individual studies included in the evidence base. Study risk of bias (sometimes referred to as threats to internal validity) is a concern because it may affect the direction or magnitude of the study's observed effect. It encompasses biases in participant selection, study performance, attrition, and outcome detection (Higgins et al. 2011). RCTs are considered the gold standard for establishing the efficacy of an intervention because a well-designed and well-conducted RCT is expected to protect against possible selection bias through the randomization process; and performance bias through allocation concealment and masking of outcome assessors and, when possible based on the treatment, masking of participants and clinicians.

Conducting an RCT (and relatedly, establishing a body of evidence of more than one RCT) to evaluate treatment for patients with a RD may be difficult because low disease incidence may make it challenging to enroll a sufficient number of participants (Behera et al. 2007). Enrolling a sample large enough to be sufficiently powered to evaluate an outcome of interest can take years. Also, if the available small pool of patients with the rare condition varies substantially on important health or sociodemographic factors, an RCT is at greater risk (than it might otherwise be) that the comparison groups will differ in key characteristics from intervention groups at baseline. This problem further complicates the methodology for evaluating outcome comparisons or limits the applicability of the results to other patients with the disease (or both).

Table 4 above presented numerous adaptations to trial designs; these adaptations are intended to improve the possibility of conducting viable trials involving patients with RDs, even though at increased risk of bias. Systematic reviewers will need to evaluate each selected study design in relation to potential for bias and the approach that researchers have taken to protect against that bias. Day (2011) supports the approach of using evidence about treatment patients with a RD from more smaller trials rather than one large trial because "every clinical trial ever carried out has some degree of bias inherent in it" (Day 2011).

When trials are not available or when answering certain study questions (such as harms or rare benefit outcomes), systematic reviewers need to consider study designs other than trials. The Committee for Medicinal Products for Human Use offers a "hierarchy" of evidence (Riegman et al. 2008). The potential for risk of bias generally rises as one goes down this list of study types:

- Meta-analyses of good-quality RCTs that all show consistent results
- Individual RCTs
- Meta-analyses of observational studies
- Individual observational studies
- Published case reports
- Anecdotal case reports
- Opinions of experts in the field.

As presented in Table 4, even within these broad categories, researchers may have adapted their approaches to studying patients with RDs because of limitations in the available data. Such modifications to basic types of study designs may be dictated by numbers of participants, anticipated progression of the disease, or other characteristics of the RDs in question.

### *Directness*

Directness concerns whether the evidence links an intervention directly to a health outcome that is of interest to the review's audience and users and whether the evidence is from head-to-head treatment comparisons. Although direct evidence is preferable, indirect evidence may be all that is available; this may include results from laboratory tests, intermediate outcomes, or in some cases reports from proxy respondents.

When a disease is rare and the mechanism of action is not well understood, studying indirect evidence may be useful. More specifically, focusing on key intermediate steps along the causal pathway, with the eventual goal of linking together more than one body of evidence, can produce important information for patients, clinicians, and policymakers. Using mechanistic evidence, rarely considered in strength of evidence grading schemes, is one approach to obtaining helpful indirect data (Goodman and Gerson 2013). This technique focuses on identifying and understanding the mechanism of action of an intervention, referred to as a target biologic mechanism, a single intermediate step between the intervention and the outcome. These data can be particularly useful in developing propensity scores and carrying out Bayesian analyses. A conceptual framework for considering mechanistic evidence is discussed in greater detail below.

### *Consistency*

Consistency concerns the degree to which included studies find the same direction or magnitude of effect for a particular outcome. For a body of evidence to be considered consistent requires, at a minimum, the inclusion of two or more studies.

### *Precision*

Precision involves the evaluation of the degree of certainty around an effect estimate for a particular outcome. Precision is based on the results from a meta-analysis or, if a meta-analysis is not possible, on the narrowness of the range of confidence intervals from the included studies. Because precision is related to the notion that the body of evidence is adequately powered (the optimal information size is met), investigations of patients with RDs may be particularly prone to imprecision when evidence is (often) limited to a small number of small studies.

### *Reporting Bias*

Reporting bias includes publication bias (nonreporting of an entire study), selective outcome reporting bias (nonreporting of planned outcomes), and selective analysis reporting (such as manipulation of cutpoints to support study goals). Evaluating reporting bias requires reviewers to compare a study as proposed and that same study as reported at completion. Therefore, this domain for grading SOE is typically limited to study designs that include a protocol, commonly only RCTs (Berkman et al. 2014).

### *Additional Domains*

Three optional domains in the AHRQ SOE approach may be of particular relevance in evaluating evidence of treatment effectiveness in RDs: dose-response association, plausible confounding that would decrease an observed effect, and magnitude of effect. For example, if the evidence is based on a small body of literature or limited to non-trials, reviewers might be able to upgrade the SOE grade for that evidence if the effect is dramatic and large. Behera and colleagues suggest a 10-times rule; when differences between treatment options exceed such a large threshold, evidence of a treatment effect is more reliable than evidence with lower levels of such differences, even if the study design is not a trial and has an increased risk for bias (Behera et al. 2007).

## Considerations for Using Mechanistic Evidence

Rather than trying only to observe an outcome from an intervention, the goals of a mechanistic evaluation are to understand the mechanism of action and to integrate that understanding into the evaluation of the observational evidence, including any effect modification (Goodman and Gerson 2013). A proposed framework for evaluating mechanistic evidence creates a formal language and structure for integrating knowledge of how the intervention works into the evaluation process. The focus is on a "target," a necessary step along a sufficient path in the

causal pathway between the intervention and the outcome. A closely related concept is the prior probability distribution functions in Bayesian approaches. In a mechanistic evidence approach, the measure of a target effect is a biomarker.

The framework for evaluating the strength of mechanistic evidence considers the intervention's target effect in nonhuman models, the clinical impact of the target effect in nonhuman models, the predictive power of a nonhuman model for an effect in humans, the predictive power of the target effect model, the predictive power of the clinical effect model, the intervention's target effect in human disease states, and the clinical impact of the target effect in human disease states.

## Conclusion

RDs are challenging for the patients who live with them, the physicians who diagnose and treat them, and the researchers who study them. We reviewed three approaches to addressing the challenges in studying RDs: registries, which make patients easier to locate and recruit, and provide efficient collection of standard data for analyses and monitoring; biobanks, which allow investigation of biomarkers without primary recruitment of patients; and study designs that are optimal for studies of the effectiveness of RD therapies.

Over the past several years, major advances have been made in developing RD patient registries and conducting RD research. Although inadequate data standardization and harmonization continues to present challenges to linking data across registries, new open-source registry platforms and common data elements provide the infrastructure needed to allow greater standardization. The development of virtual biobanks and of best practices for the management and governance of physical biobanks have increased the value of even small collections or small samples of biospecimens. New methodological research has resulted in study designs tailored for RDs or small populations, and reaching valid conclusions based on small bodies of evidence.

Our review did identify several areas that need further research. Most pressing may be the need to integrate policies and procedures for RD registries with best practices about designing and conducting studies and study design and grading strength of evidence. Similarly, identifying the types of analyses needed to answer important research questions and selecting the most robust and defensible methods are also critically important. Methodological research is needed to develop improved methods for the evaluation of the representativeness of RD registries that solicit participation by appeals on the Internet or from advocacy groups, and to investigate the validity of using registries to evaluate side effects and effectiveness of therapeutics after their approval for clinical use.

RD researchers and registry developers need to consider existing and new approaches to study and registry design to maximize the information gained from their RD registries and research.

## References


AlAama, J., T. D. Smith, A. Lo, H. Howard, A. A. Kline, M. Lange, J. Kaput, and R. G. Cotton. 2011. Initiating a Human Variome Project Country Node. *Hum Mutat* 32 (5):501-6.

Allen, K. D., E. J. Kasarskis, R. S. Bedlack, M. P. Rozear, J. C. Morgenlander, A. Sabet, L. Sams, J. H. Lindquist, M. L. Harrelson, C. J. Coffman, and E. Z. Oddone. 2008. The National Registry of Veterans with Amyotrophic Lateral Sclerosis. *Neuroepidemiology* 30 (3):180-190.

Andrews, J., G. Guyatt, A. D. Oxman, P. Alderson, P. Dahm, Y. Falck-Ytter, M. Nasser, J. Meerpohl, P. N. Post, R. Kunz, J. Brozek, G. Vist, D. Rind, E. A. Akl, and H. J. Schunemann. 2013. GRADE guidelines: 14. Going from evidence to recommendations: the significance and presentation of recommendations. *J Clin Epidemiol* 66 (7):719-25.

Atkins, D., K. Fink, J. Slutsky, Agency for Healthcare Research and Quality, and North American Evidence-based Practice Centers. 2005. Better information for better health care: the Evidence-based Practice Center program and the Agency for Healthcare Research and Quality. *Ann Intern Med* 142 (12 Pt 2):1035-41.

Australian Government, Department of Health, Therapeutic Goods Administration. Applications for orphan drug designation. www.australia.gov.au, October 9 2014 (accessed April 13, 2015.

Aymé S, Kole A , Rodwell C 2011. RDTF Report on Patient registries the field of rare diseases: Overview of the issues surrounding the establishment, governance and financing of academic registries.

Behera, M., A. Kumar, H. P. Soares, L. Sokol, and B. Djulbegovic. 2007. Evidence-based medicine for rare diseases: implications for data interpretation and clinical trial design. *Cancer Control* 14 (2):160-6.

———. 2007. Evidence-based Medicine for rare Diseases: Implications for Data Interpretation and Clinical Trial Design. *Cancer Control* 14 (2):160-166.

Bell, Stuart A., and Catrin Tudur Smith. 2014. A comparison of interventional clinical trials in rare versus non-rare diseases: an analysis of ClinicalTrials.gov. *Orphanet Journal of Rare Diseases* 9 (1):170.

Bellgard, Matthew, Christophe Beroud, Kay Parkinson, Tess Harris, Segolene Ayme, Gareth Baynam, Tarun Weeramanthri, Hugh Dawkins, and Adam Hunter. 2013. Dispelling myths about rare disease registry system development. *Source Code for Biology and Medicine* 8:21-21.

Bellgard, Matthew I., Andrew Macgregor, Fred Janon, Adam Harvey, Peter O'Leary, Adam Hunter, and Hugh Dawkins. 2012. A modular approach to disease registry design: Successful adoption of an internet-based rare disease registry. *Human Mutation* 33 (10):E2356-E2366.

Bellgard, Matthew I., Lee Render, Maciej Radochonski, and Adam Hunter. 2014. Second generation registry framework. *Source Code for Biology and Medicine* 9:14-14.

Berkman, N. D., K. N. Lohr, M. T. Ansari, E. M. Balk, R. Kane, M. McDonagh, S. C. Morton, M. Viswanathan, E. B. Bass, M. Butler, G. Gartlehner, L. Hartling, M. McPheeters, L. C. Morgan, J. Reston, P. Sista, E. Whitlock, and S. Chang. 2014. Grading the strength of a body of evidence when assessing health care interventions: an EPC update. *J Clin Epidemiol*.

Beskow, L. M., and E. Dean. 2008. Informed consent for biorepositories: assessing prospective participants' understanding and opinions. *Cancer Epidemiol Biomarkers Prev* 17 (6):1440-51.

Bogaerts, J., M. R. Sydes, N. Keat, A. McConnell, A. Benson, and A. Ho. 2015. Clinical trial designs for rare diseases: studies developed and discussed by the International Rare Cancers Initiative. *European Journal of Cancer and Clinical Oncology* 51 (3):271-281.

Brooke, E.M. 1974. The current and future use of registers in health information systems. Geneva: World Health Organization.

Cambon-Thomsen, A., E. Rial-Sebbag, and B. M. Knoppers. 2007. Trends in ethical and legal frameworks for the use of human biobanks. *Eur Respir J* 30 (2):373-82.

Canadian Institutes of Health Research, Ethical Conduct for Research Involving Humans. TCPS 2 - Chapter 3. The Consent Process. Center of Genomics and Policy, 2-13-15 2014. http://www.pre.ethics.gc.ca/eng/policy-politique/initiatives/tcps2-eptc2/chapter3-chapitre3/#toc03-1 (accessed April 15, 2015.

Centers for Disease Control and Prevention. 2003. HIPAA privacy rule and public health. Guidance from CDC and the U.S. Department of Health and Human Services. *MMWR Morb Mortal Wkly Rep* 52 Suppl:1-17, 19-20.

Choquet, R., M. Maaroufi, A. de Carrara, C. Messiaen, E. Luigi, and P. Landais. 2015. A methodology for a minimum data set for rare diseases to support national centers of excellence for healthcare and research. *J Am Med Inform Assoc* 22 (1):76-85.

Ciafaloni, E., D. J. Fox, S. Pandya, C. P. Westfield, S. Puzhankara, P. A. Romitti, K. D. Mathews, T. M. Miller, D. J. Matthews, L. A. Miller, C. Cunniff, C. M. Druschel, and R. T. Moxley. 2009. Delayed diagnosis in duchenne muscular dystrophy: data from the Muscular Dystrophy Surveillance, Tracking, and Research Network (MD STARnet). *J Pediatr* 155 (3):380-5.

Cochrane Collaboration. A Cochrane Risk Of Bias Assessment Tool: for Non-Randomized Studies of Interventions (ACROBAT-NRSI). . The Cochrane Collaboration 2011. https://sites.google.com/site/riskofbiastool/ (accessed December 15, 2014.

Collis, D. J., and M.G. Rukstad. 2008. Can you say what your strategy is? *Harvard Business Review* 86 (4).

Cornu, C., B. Kassai, R. Fisch, C. Chiron, C. Alberti, R. Guerrini, A. Rosati, G. Pons, H. Tiddens, S. Chabaud, D. Caudri, C. Ballot, P. Kurbatova, A. C. Castellan, A. Bajard, P. Nony, Cresim, CRESim Project Groups Epi, L. Aarons, A. Bajard, C. Ballot, Y. Bertrand, F. Bretz, D. Caudri, C. Castellan, S. Chabaud, C. Cornu, F. Dufour, C. Dunger-Baldauf, J. M. Dupont, R. Fisch, R. Guerrini, V. Jullien, B. Kassai, P. Nony, K. Ogungbenro, D. Perol, G. Pons, H. Tiddens, A. Rosati, C. Alberti, C. Chiron, P. Kurbatova, and R. Nabbout. 2013. Experimental designs for small randomised clinical trials: an algorithm for choice. *Orphanet J Rare Dis* 8:48.

Crowther, Mark. 2013. Phase 4 research: what happens when the rubber meets the road? *ASH Education Program Book* 2013 (1):15-18.

Cui, Y., and J. Han. 2015. A proposed definition of rare diseases for China: from the perspective of return on investment in new orphan drugs. *Orphanet J Rare Dis* 10 (1):28.

Day, Simon. 2011. Evidence-Based Medicine and Rare Diseases. In *Rare Diseases Epidemiology*.

De Souza, Y. G., and J. S. Greenspan. 2013. Biobanking past, present and future: responsibilities and benefits. *AIDS* 27 (3):303-12.

DeRenzo, E. G., and J. Moss. 2006. *Writing Clinical Research Protocols*. Edited by E. G. DeRenzo and J. Moss. London: Academic Press.

Eder, J., H. Gottweis, and K. Zatloukal. 2012. IT Solutions for Privacy Protection in Biobanking. *Public Health Genom* 15 (5):254-262.

Ford, J. G., M. W. Howerton, S. Bolen, T. L. Gary, G. Y. Lai, J. Tilburt, M. C. Gibbons, C. Baffi, R. F. Wilson, C. J. Feuerstein, P. Tanpitukpongse, N. R. Powe, and E. B. Bass. 2005. Knowledge and access to information on recruitment of underrepresented populations to cancer clinical trials. *Evid Rep Technol Assess (Summ)* (122):1-11.

Forrest, C. B., R. J. Bartek, Y. Rubinstein, and S. C. Groft. 2011. The case for a global rare-diseases registry. *Lancet* 377 (9771):1057-9.

Fung KW, Richesson R, Bodenreider O. 2014. Coverage of Rare Disease Names in Standard Terminologies and Implications for Patients, Providers, and Research. *AMIA Annu Symp Proc*:564-570.

Gagne, J. J., L. Thompson, K. O'Keefe, and A. S. Kesselheim. 2014. Innovative research methods for studying treatments for rare diseases: Methodological review. *BMJ Case Reports* 349:g6802.

Gerss, J. W., and W. Kopcke. 2010. Clinical trials and rare diseases. *Adv Exp Med Biol* 686:173-90.

Gillio-Meina, C., G. Cepinskas, E. L. Cecchini, and D. D. Fraser. 2013. Translational research in pediatrics II: blood collection, processing, shipping, and storage. *Pediatrics* 131 (4):754-66.

Gliklich, R., N. A. Dreyer, M. B. Leavy, P. Velentgas, and L. Khurana. 2012. Standards in the Conduct of Registry Studies for Patient-Centered Outcomes Research. edited by P.-C. O. R. Institute. Washington, DC.

Gliklich, R., N. Dreyer, and M. Leavy, eds. 2014. *Registries for Evaluating Patient Outcomes: A User's Guide*. Third ed. 2 vols. Rockville, MD: Agency for Healthcare Research and Quality.

Global Rare Diseases Patient Registries Data Repository (The NIH/NCATS GRDR® Program) https://grdr.ncats.nih.gov/index.php?option=com_content&view=article&id=3&Itemid=5 (accessed June 17, 2015).

Goodman, S. N., and J. Gerson. 2013. Mechanistic Evidence in Evidence-Based Medicine: A Conceptual Framework. Rockville, MD: AHRQ Methods for Effective Health Care.

———. 2013. Mechanistic Evidence in Evidence-based Medicine: A Conceptual Framework. In *AHRQ Methods for Effective Health Care*. Rockville, MD: Agency for Healthcare Research and Quality.

Gottweis, H., F. Bignami, E. Rial-Sebbag, R. Lattanzi, and M. Macek, Jr. 2012. Biobanks for Europe A challenge for governance - Report of the Expert Group on Dealing with Ethical and Regulatory Challenges of International Biobank Research Luxembourg.

Green, R. C., J. S. Berg, W. W. Grody, S. S. Kalia, B. R. Korf, C. L. Martin, A. L. McGuire, R. L. Nussbaum, J. M. O'Daniel, K. E. Ormond, H. L. Rehm, M. S. Watson, M. S. Williams, L. G. Biesecker, Genetics American College of Medical, and Genomics. 2013. ACMG recommendations for reporting of incidental findings in clinical exome and genome sequencing. *Genet Med* 15 (7):565-74.

Grinnon, Stacie T., Kristy Miller, John R. Marler, Yun Lu, Alexandra Stout, Joanne Odenkirchen, and Selma Kunitz. 2012. NINDS Common data element project – approach and methods. *Clinical trials (London, England)* 9 (3):322-329.

Gupta, S., M. E. Faughnan, G. A. Tomlinson, and A. M. Bayoumi. 2011. A framework for applying unfamiliar trial designs in studies of rare diseases. *J Clin Epidemiol* 64 (10):1085-94.

Guyatt, G., A. D. Oxman, E. A. Akl, R. Kunz, G. Vist, J. Brozek, S. Norris, Y. Falck-Ytter, P. Glasziou, H. DeBeer, R. Jaeschke, D. Rind, J. Meerpohl, P. Dahm, and H. J. Schunemann. 2011. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. *J Clin Epidemiol* 64 (4):383-94.

Higgins, J. P., D. G. Altman, P. C. Gotzsche, P. Juni, D. Moher, A. D. Oxman, J. Savovic, K. F. Schulz, L. Weeks, J. A. Sterne, Group Cochrane Bias Methods, and Group Cochrane Statistical Methods. 2011. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ* 343:d5928.

Higgins, J. P. T., D. G. Altman, and J. A. C. Sterne. 2011. Chapter 8: Assessing risk of bias in included studies. In *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 (updated March 2011)*, ed. J. P. T. Higgins and S. Green: The Cochrane Collaboration.

Hilbert, JamesE, Tetsuo Ashizawa, JohnW Day, ElizabethA Luebbe, WilliamB Martens, MichaelP McDermott, Rabi Tawil, CharlesA Thornton, and RichardT Moxley, III. 2013. Diagnostic odyssey of patients with myotonic dystrophy. *Journal of Neurology* 260 (10):2497-2504.

Holland, Nina T., Martyn T. Smith, Brenda Eskenazi, and Maria Bastaki. 2003. Biological sample collection and processing for molecular epidemiological studies. *Mutation Research/Reviews in Mutation Research* 543 (3):217-234.

Horton, D., P. Mehta, and V. C. Antao. 2014. Quantifying a nonnotifiable disease in the united states: The national amyotrophic lateral sclerosis registry model. *JAMA* 312 (11):1097-1098.

Institute of Medicine. Committee on Strategies for Responsible Sharing of Clinical Trial Data. Institute of Medicine, May 5, 2014 2014. http://www.iom.edu/activities/research/sharingclinicaltrialdata.aspx (accessed May 7, 2015.

———, ed. 2001. *Small clinical trials: issues and challenges*. Edited by C. H. Evans and S. T. Ildstad. Washington, DC: National Academies Press.

Jarvik, G. P., L. M. Amendola, J. S. Berg, K. Brothers, E. W. Clayton, W. Chung, B. J. Evans, J. P. Evans, S. M. Fullerton, C. J. Gallego, N. A. Garrison, S. W. Gray, I. A. Holm, I. J. Kullo, L. S. Lehmann, C. McCarty, C. A. Prows, H. L. Rehm, R. R. Sharp, J. Salama, S. Sanderson, S. L. Van Driest, M. S. Williams, S. M. Wolf, W. A. Wolf, Merge Act- R. O. R. Committee e, Cerc Committee, Cser Act-ROR Working Group, and W. Burke. 2014. Return of genomic results to research participants: the floor, the ceiling, and the choices in between. *Am J Hum Genet* 94 (6):818-26.

Johnson, Kimberly J., Nancy L. Mueller, Katherine Williams, and David H. Gutmann. 2014. Evaluation of participant recruitment methods to a rare disease online registry. *American Journal of Medical Genetics Part A* 164 (7):1686-1694.

Kaye, Wendy E., Marchelle Sanchez, and Jennifer Wu. 2014. Feasibility of creating a National ALS Registry using administrative data in the United States. *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration* 15 (5-6):433-439.

Korn, Edward L., Lisa M. McShane, and Boris Freidlin. 2013. Statistical Challenges in the Evaluation of Treatments for Small Patient Populations. *Science Translational Medicine* 5 (178):178sr3.

Lo, B. 2015. Sharing clinical trial data: maximizing benefits, minimizing risk. *Journal of the American Medical Association* 313 (8):793-4.

Lowrance, W. W. 2012. *Privacy, Confidentiality, and Health Research*. Cambridge: Cambridge University Press.

Malek, Angela M., David E. Stickler, Vinicius C. Antao, and D. Kevin Horton. 2014. The National ALS Registry: A recruitment tool for research. *Muscle & Nerve* 50 (5):830-834.

Mascalzoni, D., E. S. Dove, Y. Rubinstein, H. J. S. Dawkins, A. Kole, and P. McCormack. 2014. International Charter of Principles for Sharing Bio-specimens and Data. *European Journal of Human Genetics*.

Maschke, K. J. 2008. Biobanks: DNA and research. In *From Birth to Death and Bench to Clinic: The Hastings Center Bioethics Briefing Book forJournalists, Policymakers, and Campaigns*, ed. M. Crowley. Garrison, NY: The Hastings Center.

Mathews, K. D., C. Cunniff, J. R. Kantamneni, E. Ciafaloni, T. Miller, D. Matthews, V. Cwik, C. Druschel, L. Miller, F. J. Meaney, J. Sladky, and P. A. Romitti. 2010. Muscular Dystrophy Surveillance Tracking and Research Network (MD STARnet): case definition in surveillance for childhood-onset Duchenne/Becker muscular dystrophy. *J Child Neurol* 25 (9):1098-102.

McCarty, C. A., R. L. Chisholm, C. G. Chute, I. J. Kullo, G. P. Jarvik, E. B. Larson, R. Li, D. R. Masys, M. D. Ritchie, D. M. Roden, J. P. Struewing, W. A. Wolf, and Merge Team e. 2011. The eMERGE Network: a consortium of biorepositories linked to electronic medical records data for conducting genomic studies. *BMC Med Genomics* 4:13.

Miller, L. A., P. A. Romitti, C. Cunniff, C. Druschel, K. D. Mathews, F. J. Meaney, D. Matthews, J. Kantamneni, Z. F. Feng, N. Zemblidge, T. M. Miller, J. Andrews, D. Fox, E. Ciafaloni, S. Pandya, A. Montgomery, and A. Kenneson. 2006. The muscular Dystrophy Surveillance Tracking and Research Network (MD STARnet): surveillance methodology. *Birth Defects Res A Clin Mol Teratol* 76 (11):793-7.

Montserrat Moliner, A., and J. Waligora. 2013. The European union policy in the field of rare diseases. *Public Health Genomics* 16 (6):268-77.

Moxley, R. T., 3rd, S. Pandya, E. Ciafaloni, D. J. Fox, and K. Campbell. 2010. Change in natural history of Duchenne muscular dystrophy with long-term corticosteroid treatment: implications for management. *J Child Neurol* 25 (9):1116-29.

National Cancer Institute. NCI Best Practices for Biospecimen Resources. Office of Biorepositories and Biospecimen Research, National Cancer Institute, National Institutes of Health, U.S. Department of Health and Human Services 2011. http://biospecimens.cancer.gov/bestpractices/2011-NCIBestPractices.pdf (accessed April 13, 2015.

National Institutes of Health. 2012. Protecting Personal Health Information in Research: Understanding the HIPAA Privacy Rule. Washington, DC: Department of Health and Human Services.

———. NIH Sharing Policies and Related Guidance on NIH-Funded Research Resources. National Institutes of Health, February 26 2015. http://grants.nih.gov/grants/sharing.htm (accessed April 15, 2015.

National Organization for Rare Disorders (NORD), Rare Diseases Europe (EURODIS), and Canadian Organization for Rare Disorders (CORD). 2012. Joint Declaration of 10 Key Principles for Rare Disease Patient Registries.

Nigwekar, SagarU, CraigA Solid, Elizabeth Ankers, Rajeev Malhotra, William Eggert, Alexander Turchin, Ravil Thadhani, and CharlesA Herzog. 2014. Quantifying a Rare Disease in Administrative Data: The Example of Calciphylaxis. *Journal of General Internal Medicine* 29 (3):724-731.

NINDS Common Data Elements. Project Overview. NINDS, NIH, USA.gov, March 26 2015. http://www.commondataelements.ninds.nih.gov/ProjReview.aspx#tab=Benefits_of_CDEs (accessed March 23, 2015.

O'Brien, Stephen J. 2009. Stewardship of Human Biospecimens, DNA, Genotype, and Clinical Data in the GWAS Era*. *Annual Review of Genomics and Human Genetics* 10 (1):193-209.

Orphanet. The Portal for Rare Diseases and Orphan Drugs. Orpahnet/INSERM US14, April 12 2014. http://www.orpha.net/consor/cgi-bin/ResearchTrials_RegistriesMaterials.php?lng=EN (accessed April 13, 2015.

Otlowski, M., D. Nicol, and M. Stranger. Biobanks Information Paper. Australian Government, National Health and Medical Research Council 2010. https://www.nhmrc.gov.au/_files_nhmrc/file/your_health/egenetics/practioners/biobanks_information_paper.pdf (accessed April 15, 2015.

Overhage, J. M., P. B. Ryan, C. G. Reich, A. G. Hartzema, and P. E. Stang. 2012. Validation of a common data model for active safety surveillance research. *J Am Med Inform Assoc* 19 (1):54-60.

Parodi, B. 2015. Biobanks: A Definition. In *Ethics, Law and Governance of Biobanking*, ed. D. Mascalzoni: Springer.

Quon, Bradley S, and Christopher H Goss. 2011. A story of success: continuous quality improvement in cystic fibrosis care in the USA. *Thorax* 66 (12):1106-1108.

*Rare Diseases Act of 2002*. 107-280. 107th, November 6.

Relevo, R., and H. Balshem. 2011. Finding evidence for comparing medical interventions: AHRQ and the Effective Health Care Program. *Journal of Clinical Epidemiology* 64 (11):1168-1177.

Richesson, Rachel, and Kendra Vehik. 2010. Patient registries: utility, validity and inference. In *Advances in experimental medicine and biology*,87-104.

Riegman, P. H., M. M. Morente, F. Betsou, P. de Blasio, P. Geary, and Marble Arch International Working Group on Biobanking for Biomedical Research. 2008. Biobanking for better healthcare. *Mol Oncol* 2 (3):213-22.

Rubinstein, YR, McInnes P. NIH/NCATS/GRDR® Common Data Elements: A leading force for standardized data collection. 2015. *Contemp Clin Trials* 42:78-80.

Rubinstein, YR, B Karp, N Lockhart, C Grady, and SC Groft. 2014. Informed consent template for patient participation in rare disease registries linked to biorepositories. *Rare Diseases and Orphan Drugs* 1 (2):69-75.

Schunemann, H. J., A. D. Oxman, J. Brozek, P. Glasziou, R. Jaeschke, G. E. Vist, J. W. Williams, Jr., R. Kunz, J. Craig, V. M. Montori, P. Bossuyt, G. H. Guyatt, and Grade Working Group. 2008. Grading quality of evidence and strength of recommendations for diagnostic tests and strategies. *BMJ* 336 (7653):1106-10.

Sengupta, S., N. S. Calman, and G. Hripcsak. 2008. A model for expanded public health reporting in the context of HIPAA. *J Am Med Inform Assoc* 15 (5):569-74.

Shabihkhani, Maryam, Gregory M. Lucey, Bowen Wei, Sergey Mareninov, Jerry J. Lou, Harry V. Vinters, Elyse J. Singer, Timothy F. Cloughesy, and William H. Yong. 2014. The procurement, storage, and quality assurance of frozen blood and tissue biospecimens in pathology, biorepository, and biobank settings. *Clinical Biochemistry* 47 (4–5):258-266.

Sharkey, E. K., N. L. Zoellner, S. Abadin, D. H. Gutmann, and K. J. Johnson. 2014. Validity of participant-reported diagnoses in an online patient registry: A report from the NF1 Patient Registry Initiative. *Contemp Clin Trials* 40C:212-217.

Singh, S., S. M. Chang, D. B. Matchar, and E. B. Bass. 2012. Grading a body of evidence on diagnostic tests. In *Methods Guide for Medical Test Reviews [Internet]*, ed. S. M. Chang, D. B. Matchar, G. W. Smetana and C. A. Umscheid. Rockville, MD: Agency for Healthcare Research and Quality.

Smith, Maureen E., and Sharon Aufox. 2013. Biobanking: The Melding of Research with Clinical Care. *Current genetic medicine reports* 1 (2):122-128.

Song, P., J. Gao, Y. Inagaki, N. Kokudo, and W. Tang. 2012. Rare diseases, orphan drugs, and their regulation in Asia: Current status and future perspectives. *Intractable Rare Dis Res* 1 (1):3-9.

Taruscio, D., E. Mollo, S. Gainotti, M. Posada de la Paz, F. Bianchi, and L. Vittozzi. 2014. The EPIRARE proposal of a set of indicators and common data elements for the European platform for rare disease registration. *Arch Public Health* 72 (1):35.

Tudur Smith, C., P. R. Williamson, and M. W. Beresford. 2014. Methodology of clinical trials for rare diseases. *Best Pract Res Clin Rheumatol* 28 (2):247-62.

U.S. National Library of Medicine. Common Data Element (CDE) Resource Portal. U.S. National Library of Medicine, January 3 2013. http://www.nlm.nih.gov/cde/ (accessed May 7, 2015.

van der Lee, J. H., J. Wesseling, M. W. Tanck, and M. Offringa. 2010. Sequential design with boundaries approach in pediatric intervention research reduces sample size. *J Clin Epidemiol* 63 (1):19-27.

Vaught, J., J. Rogers, T. Carolin, and C. Compton. 2011. Biobankonomics: developing a sustainable business model approach for the formation of a human tissue biobank. *J Natl Cancer Inst Monogr* 2011 (42):24-31.

Venance, S. L., B. E. Herr, and R. C. Griggs. 2007. Challenges in the design and conduct of therapeutic trials in channel disorders. *Neurotherapeutics* 4 (2):199-204.

Viswanathan, M., M. T. Ansari, N. D. Berkman, S. Chang, L. Hartling, and M. McPheeters. 2008. Assessing the Risk of Bias of Individual Studies in Systematic Reviews of Health Care Interventions. In *Methods Guide for Effectiveness and Comparative Effectiveness Reviews*. Rockville (MD).

Wolfe, F., and K. Michaud. 2010. The Hawthorne effect, sponsored trials, and the overestimation of treatment effectiveness. *J Rheumatol* 37 (11):2216-20.

Wong, Siaw H., Belinda J. McClaren, Alison Dalton Archibald, Alice Weeks, Tess Langmaid, Monique M. Ryan, Andrew Kornberg, and Sylvia A. Metcalfe. 2015. A mixed methods study of age at diagnosis and diagnostic odyssey for Duchenne muscular dystrophy. *Eur J Hum Genet*.

Workman, T. A. 2013. Engaging Patients in Information Sharing and Data Collection: The Role of Patient-Powered Registries and Research Networks [Internet]. Rockville, MD: Agency for Healthcare Research and Quality.

# Appendix A: Methods for Landscape Review of Rare Disease Registries and Stewardship of Biospecimens and Biobanks

Purpose: Per PCORI's guidance, focus is on best practices for designing new registries for research.

Sources: (1) Personal knowledge of key references and investigators; (2) Pub Med; (3) material on websites of NIH and AHRQ; (4) previously unidentified, but relevant, references cited in articles that were identified in Sources # 1-3

Search Strategies

1. Pub Med
   i. Criteria:
      - Published in English;
      - Published in 2005 or more recently, unless reference was unique or especially outstanding, but published before 2005; and
      - Contained content that did not duplicate more recently published references.
   ii. Conducted a search for the following four combination of terms:
      - "Registry" AND "Rare Disease"
      - "biospecimen" AND "rare disease";
      - "biobank" AND "rare disease";
      - "stewardship" AND "biospecimen OR biobanks" AND "rare disease"
   iii. For references that appeared relevant based on their abstracts, examined "Related Citations" identified by Pub Med and relevant references that cited the reference under consideration.
2. NIH and AHRQ websites: searched the terms listed above for Pub Med.
3. Read key articles and materials from NIH and ARHQ websites and identified references that were cited, but had not been identified by the Pub Med search.
4. Sources recommended by PCORI Working Group.
5. Judged that search was complete when methods failed to identify new relevant references.

**Inclusion and exclusion criteria**

| Criterion | Include | Exclude |
|---|---|---|
| Populations | Studies of research registries for rare diseases<br>[Studies on registries for specific RD communities considered as resources permit.]<br><br>For Stewardship of biospecimens and biobanks, focused on references that addressed issues likely to occur in the United States, because of our laws or systems of health care or health insurance. | Reports of registries for non-rare diseases, *except* for comprehensive reports (e.g., Gliklich, 2014) |
| Interventions | None specified | None specified |
| Comparator | None specified | None specified |
| Outcomes | None specified | None specified |
| Timing | Studies published after 2004, except unique or very informative studies published earlier. | Published before 2005 unless unique or very informative. |
| Setting | None specified (preference for U.S.) | None specified |
| Language | English | Non-English |

# Appendix B: Methods for Landscape Review for Issues About Study Design and Strength of Evidence for Rare Disease Research

**Key Questions**

1. What are study designs that can be used to evaluate therapies for rare diseases? What are their applications and constraints?

2. What are strength of evidence systems that can be used for evaluating therapies for rare diseases? What are their applications and constraints?

**Sources**

1. PubMed

2. SRC Methods Library (curated database of methodological references from PubMed, Cochrane, AHRQ, and several other sources)

2. JGIM special edition on rare diseases

**Search strings**

**Search in PubMed for strength of evidence systems (2/3/2015)**

| Search | Query | Items found |
|---|---|---|
| #8 | Search (#7 and #3) | 84 |
| #7 | Search (("Practice Guideline" [Publication Type]) OR "Practice Guidelines as Topic"[Mesh]) OR "Evidence-Based Medicine"[Mesh] or “strength of evidence” | 146352 |
| #3 | Search "Rare Diseases"[Mesh] | 5285 |

- AND in builder
- OR in builder
- NOT in builder
- Delete from history
- Show search results
- Show search details
- AND in builder
- OR in builder
- NOT in builder
- Delete from history
- Show search results
- Show search details
- Save in My NCBI

- AND in builder
- OR in builder
- NOT in builder
- Show search results
- Save as a My NCBI Collection

**Search in PubMed for study design issues (2/3/2015)**

| Search | Query | Items found |
|---|---|---|
| #17 | Search (#16 and #3) | 166 |
| #16 | Search ("Cross-Over Studies"[Mesh] OR "Non-Randomized Controlled Trials as Topic"[Mesh]) OR ( "Controlled Clinical Trials as Topic"[Mesh] OR "Pragmatic Clinical Trials as Topic"[Mesh] OR "Clinical Trials as Topic"[Mesh] OR "Compassionate Use Trials"[Mesh] OR "Clinical Trials, Phase III as Topic"[Mesh] OR "Randomized Controlled Trials as Topic"[Mesh] OR "Clinical Trials, Phase I as Topic"[Mesh] OR "Clinical Trials, Phase II as Topic"[Mesh] OR "Clinical Trials, Phase IV as Topic"[Mesh] OR "Multicenter Studies as Topic"[Mesh] OR "Drugs, Investigational"[Mesh] OR "Therapies, Investigational"[Mesh] ) | 319738 |
| #3 | Search "Rare Diseases"[Mesh] | 5285 |

Search string in SRC Methods Library: [KW; rare] = 47 articles (2/3/2015)
Hand searches from references and experts: 10

**Inclusion and exclusion criteria**

| PICOTS | Include | Exclude |
|---|---|---|
| Populations | Studies of methods on rare diseases | Studies of methods on rare outcomes for non-rare diseases |
| Interventions | Studies on study designs or strength of evidence systems | Studies on statistical methods or outcomes without comment on design or strength of evidence |
| Comparator | None specified | None specified |
| Outcomes | None specified | None specified |
| Timing | None specified | None specified |
| Setting | None specified | None specified |
| Language | English | Non-English |