# Evaluating MDA's MOVR Data Hub as a Source for Real-World Data

Elisabeth A Kilroy, PhD[1], Jessica Waits[1], Sharon Hesterlee, PhD[1]

[1] Muscular Dystrophy Association, 161 N. Clark, Suite 3550, Chicago, IL 60601

## Abstract

**Purpose:** The number of therapies in the drug development pipeline for neuromuscular diseases (NMDs) is increasing at a rate that outcompetes patient data availability. The Muscular Dystrophy Association (MDA) created the neuroMuscular ObserVational Research Data Hub (MOVR) to address this significant data shortage and to provide a database for clinicians, drug developers, and regulators with which to make data-driven efficacy and safety decisions. This study describes the mechanisms employed by the MOVR team to evaluate whether data captured within the data hub are compliant with the recent FDA draft guidances on real-world data (RWD).

**Background:** MOVR represents the first data hub that aggregates clinical and genetic data across multiple NMDs, including ALS, BMD, DMD, FSHD, LGMD, Pompe disease, and SMA. Data are collected using electronic case report forms (eCRFs) that capture clinically relevant data for demographics, diagnosis, disease progression, and discontinuation. Last year, the FDA released draft guidances that focus on using RWD to develop real-world evidence (RWE) to support regulatory decisions. The draft guidance entitled "Assessing Registries to Support Regulatory Decision-Making for Drug and Biological Products" details what is expected if registry data are used in a submission.

**Methods:** To evaluate MOVR's compliance with this draft guidance, the MOVR team turned each criterion stated by the FDA into a question that could be answered 'Yes' or 'No' to regarding whether MOVR satisfied the criterion. Questions were transferred to a table with three columns: (1) FDA Guideline, (2) Satisfied by MOVR?, and (3) How MOVR Satisfies Guideline. With each question occupying its own row, the MOVR team then identified whether MOVR satisfies the criterion and provided detailed explanations as to how MOVR meets the criterion.

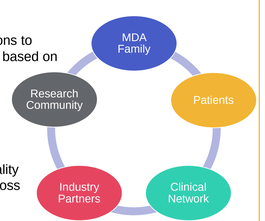
**Results:** The draft guidance mentioned several key topics, including data dictionary, rules for data validations, procedures for data collection, curation, management and storage, data access, data protection, version control, and updating eCRFs to reflect changing clinical information. MOVR meets 77 percent of these criteria (24 out of 31 criteria) and the team is currently implementing strategies to address those that were not satisfied.

**Conclusions:** As a centralized clinic-entered data hub, analyses demonstrate that MOVR serves as a rigorous platform that could become an important component of the drug development pipeline for NMDs. MDA is committed to ensuring that MOVR is compliant with the final guidances in hopes that this will help forestall the development of proprietary industry databases and siloing of patient data.

## Goals for MOVR

MOVR aims to serve the entire neuromuscular disease community as the first data hub that will aggregate clinical, genetic and patient-reported data for multiple NMDs to improve health outcomes and accelerate drug development. The goals for MOVR are:

- Gain a better understanding of the course of illness for specific NMDs
- Collect data about genotype-phenotype correlations to allow for better prediction of disease progression based on genetic information
- Collect longitudinal patient data that will allow benchmarking of best clinical practices
- Use data to develop and implement a clinical quality improvement program for MDA Care Centers across the country
- Provide outcome-related information about MDA Care Centers for families seeking medical care
- Establish a database of individuals eligible for clinical trials in NMDs to ease the burden of clinical trial recruitment and accelerate drug development




## MOVR as a Centralized Data Hub

About 10 years ago, MDA recognized that there was a significant data shortage in the NMD space and started crafting strategic approaches to accelerate data collection and its use by researchers, clinicians, and drug developers. One strategy that was identified was to leverage the MDA Care Center Network, which is comprised of over 150 care centers and 2,400 clinical providers across the United States, as a source for efficiently capturing clinical data and growing a longitudinal dataset. Specifically, each year, over 90,000 medical visits are conducted and over 60,000 individuals living with NMDs receive expert care at these centers.

Data elements captured by MOVR are functional and disease-specific outcome measures that have been identified by KOLs as important to understanding disease mechanisms, tracking disease progression, and implementing standards of care.

**Core data elements captured by the eCRFs.**

| Demographics eCRF (During Enrollment) | Diagnosis eCRF* (During Enrollment) | Encounter eCRF* (During Clinical Visits) | Discontinuation eCRF (After End of Study) |
|---|---|---|---|
| Disease Type | Age at Diagnosis | Encounter Date | Date of Discontinuation |
| Enrollment Date | Age at Symptom Onset | Height | Reason for Discontinuation |
| Gender | Clinical Diagnosis | Weight | Date of Death |
| DOB | First Symptoms | Clinical Trial Participation | Cause of Death |
| Race | Family History | Surgeries | |
| Ethnicity | Genetic Testing Results | Hospitalizations | |
| Insurance | | Medications | |
| Education | | Pulmonary Devices | |
| Employment | | Assistive Devices | |
| | | Functional Testing | |
| | | Pulmonary Tests | |
| | | Referral Types | |

* Diagnosis and Encounter eCRFs contain additional unique fields for each indication.

## Data Collection and Access

**Step 1** **MOVR team trains MDA Care Centers on data collection and entry processes.**

**MOVR Onboarding**

Site Identification and Selection → Site Initiation Visit and Training → MDA Care Center is now designated as a MOVR Site

**Data Collection and Entry**

EHR = Source Data → eCRFs

**Step 2** **MOVR houses a secure PHI data environment for data extraction, transformation, and storage.**

IQVIA

**Secure PHI Data Environment**

- **Data Extraction** — Automated in-line validation alerts
- **Data Transformation** — Data quality checks via manual queries and SAS listing reviews

**Secure De-Identified Data Environment**

- **Data De-Identification** — Ongoing privacy assessments to meet standards and create limited datasets
- **Data Analysis** — Tailor data to researcher needs prior to delivery

Continuous data hosting and auditing of source and transformed data

Limited dataset ready for analysis

**Step 3** **MOVR team receives requests from researchers at different institutions.**

**Researcher Submits Data Request**
Includes type of data, selected indications, and scientific use of data

→ **MOVR Team Reviews Request**
Determines if request falls within Data Governance Policy, and if requested data are available

→ **Dataset is Delivered**
OR
**MOVR Team Performs Analyses**
OR
**Access to MOVR Visualization and Reporting Platform is Granted**

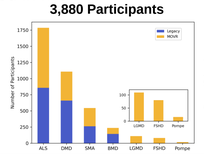
→ **Research Outputs** *(analytics, reports, publications)*

## Longitudinal Data Availability

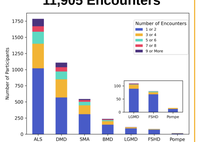
Since its inception in 2019, MOVR has experienced tremendous growth in the number of active care centers participating in the MOVR Study, the number of participants providing data to MOVR, and the number of clinical encounters captured in MOVR.

**50 Active MOVR Sites**

**3,880 Participants**




**11,905 Encounters**




**Longitudinal data availability across the 7 indications.**

**Time Between First and Most Recent Encounter[1]**

| | Average (Months) | 5 or Less Months[2] | 6 to 10 Months[2] | 11 to 25 Months[2] | 16 to 20 Months[2] | 21 to 25 Months[2] | 26 or More Months[2] |
|---|---|---|---|---|---|---|---|
| ALS | 13.7 | 315 | 271 | 167 | 117 | 74 | 153 |
| DMD | 26.5 | 20 | 170 | 99 | 71 | 87 | 309 |
| SMA | 23.4 | 25 | 87 | 42 | 33 | 25 | 128 |
| BMD | 26.0 | 8 | 23 | 24 | 20 | 9 | 54 |
| LGMD | 14.8 | 2 | 14 | 21 | 6 | 3 | 1 |
| FSHD | 14.3 | 2 | 6 | 9 | 4 | 7 | 1 |
| Pompe | 14.3 | 0 | 1 | 3 | 2 | 1 | 0 |

**Time Between Consecutive Encounters[1]**

| | Average (Months) | 5 or Less Months[3] | 6 to 10 Months[3] | 11 to 25 Months[3] | 16 to 20 Months[3] | 21 to 25 Months[3] | 26 or More Months[3] |
|---|---|---|---|---|---|---|---|
| ALS | 3.9 | 3163 | 594 | 49 | 23 | 11 | 40 |
| DMD | 7.9 | 435 | 1677 | 271 | 71 | 44 | 50 |
| SMA | 7.8 | 312 | 512 | 113 | 41 | 15 | 28 |
| BMD | 9.1 | 44 | 235 | 88 | 10 | 4 | 13 |
| LGMD | 7.2 | 34 | 40 | 21 | 2 | 2 | 0 |
| FSHD | 7.8 | 13 | 24 | 13 | 1 | 1 | 1 |
| Pompe | 8.4 | 2 | 4 | 6 | 0 | 0 | 0 |

[1] Only includes participants with at least 2 encounters
[2] Represents number of participants
[3] Represents number of encounters

Data Cutoff: 24MAY2022

## MOVR's Compliance with FDA's Proposed RWE Guidelines

The MDA launched MOVR to serve as a valuable tool for capturing a longitudinal data that could provide knowledge on disease progression for drug development as well as serve as RWD and RWE in regulatory submissions and post-approval processes. In Fall 2021, the FDA published a draft guidance entitled "Assessing Registries to Support Regulatory Decision-Making for Drug and Biological Products". This draft document provides guidelines for sponsors when selecting registries as data sources for their regulatory submissions. The MOVR team used these guidelines to assess MOVR's compliance. The below table demonstrates how MOVR satisfies the individual guidelines presented in the draft guidance.

| | | |
|---|---|---|
| 1 | **Data Dictionary** | |
| | Does the registry have an established data dictionary? | ✓ |
| | Is it made available for those who intend to use the registry data? | ✓ |
| | Does it include data elements and how the data elements are defined? | ✓ |
| | Does it include ranges and allowable values for the data elements? | ✗ |
| | Does it reference to the source data for the data elements? | ✓ |
| 2 | **Rules for Validations, Data Quality Assessments and Auditing** | |
| | Does the registry have rules for the validation of queries and edit checks of registry data? | ✓ |
| | Is it made available for those who intend to use the registry data? | ✓ |
| | Do the registry personnel and processes in place during data collection and analysis provide adequate assurance that errors are minimized, and that data integrity is sufficient? | ✓ |
| | For an electronic database, does the registry perform preventative and/or corrective actions to address changes to the data (including flagging erroneous data without deleting the erroneous data, while inserting the corrected data for subsequent use)? | ✓ |
| | For an electronic database, does the registry ensure data transferred from another data format or system are not altered in the migration process? | ✓ |
| | For an electronic database, does the registry explain auditing rules and methods used and the mitigation strategies used to reduce errors? | ✗ |
| | Does it describe the types of errors that were identified based on audit findings and how the data were corrected? | ✗ |
| | Does the registry perform routine descriptive statistical analysis to detect the extent of any missing data, inconsistent data, outliers, and losses to follow-up? | ✗ |
| 3 | **Procedures for Data Collection, Curation, Management and Storage** | |
| | Does the registry have a defined process and procedure for data collection? | ✓ |
| | Does the registry have a defined process and procedure for data curation? | ✓ |
| | Does the registry have a defined process and procedure for data management? | ✓ |
| | Does the registry have a defined process and procedure for data storage? | ✓ |
| | Does the registry have a defined process and procedure to ensure that data within the registry can be confirmed by source data? | ✗ |
| 4 | **Data Access** | |
| | Does the registry have a plan for how patients will access and interact with the registry data and the registry's data collection systems? | ✗ |
| | Does the registry have a plan for how researchers will access and interact with the registry data and the registry's data collection systems? | ✓ |
| | Does the registry have a plan for how clinicians will access and interact with the registry data and the registry's data collection systems? | ✓ |
| | Does the registry have terms and conditions for use of the registry data by parties other than the registry creator? | ✓ |
| | Does the registry conform with 21 CFR part 11, as applicable, including maintenance of access controls and audit trails to demonstrate provenance of the registry data and support traceability of the data? | ✓ |
| 5 | **Data Privacy and Security** | |
| | Does the registry adhere to applicable jurisdictional human subject protection requirements, including protecting the privacy of patient health information? | ✓ |
| | Did the registry consult with an institutional review board or independent ethics committee when developing the registry to review data collection and other procedures associated with the registry? | ✓ |
| | Does the registry have policies and procedures in place for validating the electronic systems used to collect registry data? | ✓ |
| 6 | **Version Control and Data Consistency** | |
| | For an electronic database, does the registry implement and maintain version control by documenting the date, time and originator of data entered in the registry? | ✓ |
| | For an electronic database, does the registry seek to integrate data in the registry that were previously collected using data formats or technology that are now outdated? | ✗ |
| | Are the formats and definitions of the data entered in the registry consistent over time? | ✓ |
| 7 | **Updates to Reflect Changing Clinical Information** | |
| | For an electronic database, does the registry account for changes in clinical information over time (such as criteria for disease diagnosis)? | ✓ |
| | Are changes in diagnostic criteria or clinical definitions accounted for and documented? | ✓ |

Read MDA's comments on FDA-2021-D-1146: Real-World Data: Assessing Registries to Support Regulatory Decision-Making for Drug and Biological Products – Guidance for Industry

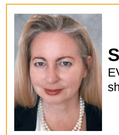
Read MDA's comments on FDA-2021-D-1214: Considerations for the Use of Real-World Data and Real-World Evidence to Support Regulatory Decision-Making for Drug and Biological Products – Guidance for Industry

## Conclusions

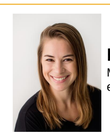
**MOVR satisfies 24 of the 31 guidelines** that would be required for the use of MOVR data in regulatory submissions according to the draft guidance. For those guidelines that MOVR does not satisfy, MDA is working diligently to develop approaches that would ensure MOVR's compliance.

**The MDA suggests that the FDA create a certification or qualification program** that registries can complete to demonstrate that they are FDA-compliant and a reputable source for RWD. This qualification program could allow for registries to prove compliance with the recommendations put forward by this guidance and the other draft guidances issued under the RWE Program without having to reassert compliance with every product submission, thus greatly reducing the resources needed for both the sponsor and the FDA. With a qualification program, the FDA can be confident in the integrity of the data being submitted, and the focus can be on the data included rather than the processes and procedures used to collect, store, and transform the data.

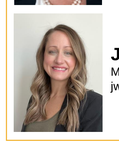
## Contact the MOVR Team



**Sharon Hesterlee, PhD**
EVP & Chief Research Officer
shesterlee@mdausa.org



**Elisabeth Kilroy, PhD**
MOVR Data Analytics
ekilroy@mdausa.org



**Jessica Waits**
MOVR Clinical Operations
jwaits@mdausa.org

**General Study Inbox**
mdamovr@mdausa.org

**Acknowledgements**
This study was conducted using data from the Muscular Dystrophy Association's neuroMuscular ObserVational Research (MOVR) Data Hub. MOVR is operated through participating MDA Care Centers with the support of participants, MOVR Site PIs, coordinators, and staff. The authors would also like to thank members of the MOVR Research Advisory Committee, including James Berry, Barry Byrne, John Day, Rod Howell, Monkol Lek, Nicholas Maragakis, and Sabrina Paganoni, for sharing their expertise in neuromuscular diseases, and the original MDA team members who spear-headed the development of the MOVR eCRFs, including Grace Pavlath, Lianna Orlando, Amanda Hayden-Phillips, and Laura Hagerty.