



Virginia Department of  
Behavioral Health &  
Developmental Services

Data Quality Monitoring Plan  
Office of Data Quality and Visualization

# Data Quality Monitoring Plan

---



## Preliminary Report

### Introduction

The Office of Data Quality and Visualization (DQV) is a program within the Office of the Chief Clinical Officer. One of the goals of DQV is to advance the use of data analytics that inform the decision making of the agency. DQV has not been able to provide this universal support due to challenges with data quality. In response, they initiated the development of a Data Quality Plan in the second quarter of SFY 2019. The ultimate goal of the Plan is to inform the establishment of a Data Quality Monitoring System.

DQV applied the clinical quality monitoring framework of Avedis Donabedian to the development of a Data Quality Monitoring Plan. Donabedian was a physician who is considered by many in the industry to be the “father of clinical quality”. In 1966 he published the pivotal paper “Evaluating the Quality of Medical Care”, providing physicians with the language and process to begin to address clinical quality. Following that article, he published many books on his framework for healthcare quality (*insert citation*).

Quality monitoring, be it for clinical care or data entry and submission, ends with remediation that relies on behavioral change driven by standards and best practices. Therefore, the use of this systematic epidemiological approach, with its foundation rooted in prevention, is appropriate. The following are the steps of quality monitoring established by Donabedian:

1. Determining what to monitor
2. Determining priorities in monitoring
3. Selecting an assessment approach
4. Formulating criteria and standards
5. Obtaining the necessary information
6. Choosing when and how to monitor
7. Constructing a monitoring system
8. Bringing about behavior change

DQV has completed the first five steps which are outlined in this document. The intent is that business owners of the data sources, with the support of DQV, will prioritize the completion of the remaining steps on their respective data sources.

## Current Efforts

### 1. Determining what to monitor

DQV chose twelve data sources as the focus for the development of this Plan. This is not an exhaustive list, but rather, a selection that vary in maturity, scale, and scope all having been identified as providing necessary information for moving towards compliance with the Settlement Agreement. Many of these sources have data accessible through the Data Warehouse, however the descriptions included in this process refer to data in the source system itself. This was a purposeful decision to draw focus to data quality monitoring at the source level. The following data sources are included in this Plan:

- Children in Nursing Facilities
- Community Consumer Submission (CCS3)
- Computerized Human Rights Information System (CHRIS) Human Rights
- Computerized Human Rights Information System (CHRIS) Serious Incident
- Independent Housing
- Individual and Family Support Program (IFSP)
- Individual Support Plan (ISP)
- Office of Licensing Information System (OLIS)
- PAIRS
- Post-Move Monitoring (PMM)
- Regional Support Teams (RST)
- Waiver Management System (WaMS)

### 2. Determining priorities in monitoring

DQV prioritized key components that either directly, or indirectly, contribute to data accuracy. Accuracy refers to how well the data reflect what they were designed to measure or otherwise represent and is one of the key expectations for data to be considered fit for use. While accuracy is only one attribute of data quality, it is a foundational attribute that if absent, should be prioritized for remediation. The focal areas of this initial step of the Plan are the following attributes of data accuracy:

#### *Unique Identifiers*

Asserting uniqueness of the entities within a data set implies that no entity exists more than once within the data set and that there is a key that can be used to uniquely access each entity (and only that specific entity) within the data set. Failure to adhere to this construct of uniqueness results in a loss of referential integrity, reliability, as well as the ability to derive accurate counts.

### *Data Validations*

Processes in place that address accuracy as data is being entered is a best practice for any data system. Front-end validations can take many forms such as edits in the software that restrict certain data types to be required in certain fields; one of the simplest examples of this is prohibiting non-numeric values in a date field.

### *Documentation*

Documentation is the primary method used to ensure that people entering, reporting, and acting on the data share a common understanding of its meaning. Without documentation, proper use of data relies on the knowledge of employees who may leave the organization at any time.

## 3. Selecting an assessment approach

Three approaches to assessing quality are structure, process, and outcome. This initial inventory focused on structural components of the data sources that effect data accuracy. Attributes of structure are more readily observable and tend to be more stable, thus an appropriate direction for this inaugural effort.

## 4. Formulating criteria and standards

In an effort to provide data stewardship to the Department, DQV has taken the lead in developing this Plan. A goal of the Plan is to provide general guidelines to the respective business owners concerning the optimum characteristics of data accuracy. These characteristics are the presence of a unique identifier, data validations, and documentation. Business owners may choose to establish more explicit criteria and standards appropriate to the unique state of their data source and this will be discussed further in the "Future Efforts" section.

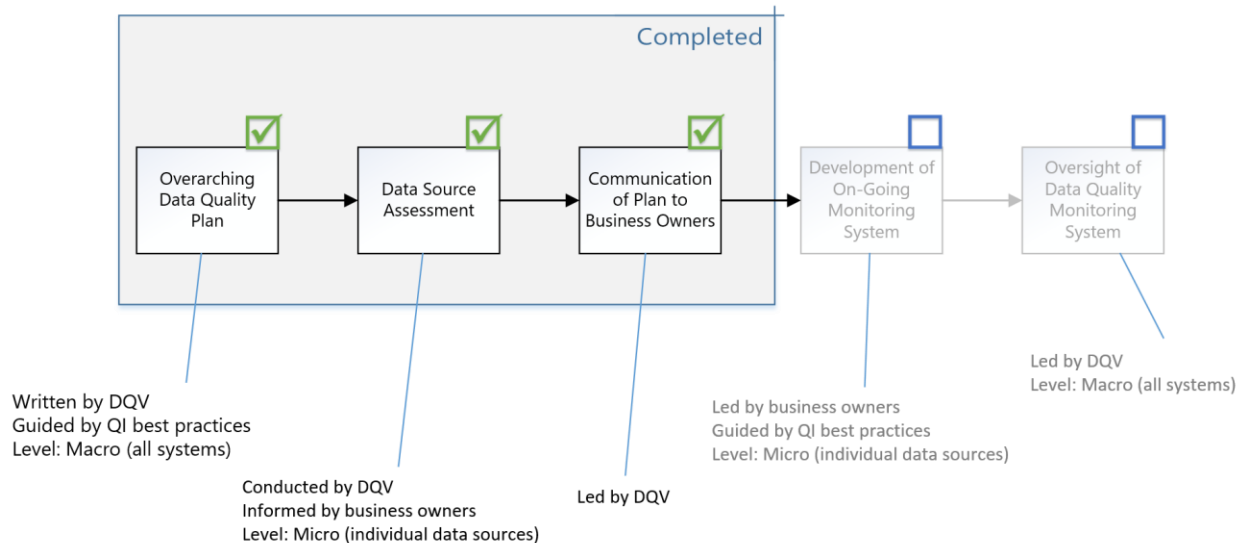
## 5. Obtaining the necessary information

The formal planning for this effort began in November 2018. Information was gathered through a combination of meetings with business owners and other subject matter experts, documentation reviews, and review of the data itself. Through this process a form was completed for each data source that includes the following information:

- Contents of the data source
- Who enters the data
- Existing data validation
- Data quality concern
- Status of the solution
- Purpose of analyzing the data
- Unique identifiers and joining
- Existing documentation
- Data quality concern - solution

## Future State

The remaining three steps guide the establishment of the ongoing quality monitoring system. Decision-making surrounding these final steps should be led by the business owners. To that end, as they develop their own quality monitoring plans, DQV would encourage them to modify details of the first five steps to better meet their ongoing monitoring needs. The remaining steps are outlined below along with a brief explanation and general recommendations by DQV.



## 6. Choosing when and how to monitor

There are three main time periods for monitoring; prospective, concurrent, and retrospective. Prospective monitoring attempts to assess and judge an action prior to its occurrence. For data quality monitoring, a prospective monitoring example is the opportunity when software for a new data source is purchased and decisions are made about how a provider will be uniquely identified.

Concurrent monitoring occurs while data is being collected. This may be as a routine process or when triggered by something like a system error. Monitoring concurrently enables timely identification and potentially, remediation however it requires process and infrastructure that may present feasibility issues.

Retrospective monitoring is the easiest to implement and maintain, and the most widely used. While this monitoring is the least timely of the three, by reviewing existing records it allows for learning to occur that can inform future data quality efforts.

There are many ways to address the question of “how to monitor”. The choice should be made by the data source owner taking into account their priorities and constraints. The choices of how to monitor will be guided by the decision that was made regarding when to monitor.

## **7. Constructing a monitoring system**

DQV recommends starting the individual source monitoring on a small scale by limiting the focus to the aforementioned aspects of data quality. It is incumbent on the business owners lead in the implementation and ongoing monitoring of their respective quality plans. DQV will provide guidance to the business owners as needed and oversight of the quality monitoring at a macro level. Initial guidance will include an evaluation of the data source using an industry-standard maturity scoring matrix.

## **8. Bringing about behavior change**

The intention of the Data Quality Plan, structural assessment, and maturity model is to provide a framework for demonstrating how data sources compare to established standards of quality and provide a path for remediation. Behavioral change around data quality includes owners of the data sources to adopt quality as a critical component to their processes.

# Components of the Assessment

The following are descriptions of the fields collected for each of the data sources. This information will comprise the “structural assessment” component of the process.

## 1. Contents of the data source

This section is a summary of the key elements within the data source. This is not meant to be an exhaustive list.

## 2. Purpose of analyzing the data

This section highlights areas of reporting that are currently being done. Data reporting varies from simple data aggregation to more complex analysis.

## 3. Who enters the data?

This section lists the groups that are responsible for entering data. This is important as data entry should be limited to only those essential to the process.

## 4. Existing data validation

This section provides a general description of existing data validations. Data validations are processes that are in place to ensure the quality of the data being entered. Data validations take many forms from as simple as having drop-down menus that restrict entry of a particular field, to conditional formatting where only certain fields are accessible based on the data entered. Details of specific data validations were not included in this initial overview. We anticipate this section growing over time as more specific details are gathered about each of the systems.

## 5. Unique identifiers and joining

This section addresses whether the dataset contains a primary key that identifies both a unique record, as well as individual and provider (where appropriate). This also includes information if there is an identifier present that would enable joining data with other sources.

## 6. Existing documentation

This section indicates the presence or absence of written documentation of the dataset. Documentation may be in the form of a data dictionary, programming code, or other informal or formal written notes. Documentation of datasets is an essential component in ensuring data quality through standard field definitions.

## 7. Data quality concern

This section will be used as the foundation of continuous quality improvement. This section lists issues that have arisen when the SME enters, stores, or reports on the data. This section includes a “solution” column that may have already been identified by the SME or that the SME and DQV came to during their interview session. The status and status notes provide insight into how far along the solution is to being implemented.