



# HIGH-QUALITY DATA DEFINITIONS

## Data Collection

1. Data is representative (crucial for equity) and is collected using clearly documented, well thought out methodologies for data collection.
2. Sources of data are dependable.
3. Data granularity is preserved by collecting data in their component forms and computing percentages, ratios, and other calculations at the state level.
4. Data collection, processing, and reporting systems are automated and data can be transmitted in an electronic, interoperable format.
5. Properly document and store data collections and results.
6. There is an intent behind data being collected. At the outset of each data collection is either a research question or CQI process. From there, state teams decide what data best satisfies this need. If data collection precedes the CQI process, often staff find that data to answer research questions or navigate the CQI process is not available.

## Data Definitions

- Establish **standard statewide definitions** and procedures for all reporting elements. **Systematic business rules** define acceptable values, character formats, and options for handling missing or unavailable data. States work with vendors to ensure that assessments and other data collection instruments align with definitions and standards. **Avoid inconsistent item response** as happens, for example, when not all data providers report the same data elements. Idiosyncratic reporting of different types of information from different sources creates gaps and errors in macro-level data aggregation.
- A **data dictionary** identifies all data elements used in collection and reporting, and describes their content and format.
- Data providers use **clearly defined, broadly understood data definitions**. All data providers must use the same definitions for the same elements. Otherwise, when the data is passed on to the state level, it will not be aggregated appropriately and will lead to inaccurate results.
- Hardware and software, along with staff training, **are configured around standard data definitions and business rules**.
- **Avoid Inconsistency over time**. Avoid having the same data element calculated, defined, or reported differently from year to year. *Longitudinal inconsistency creates the potential for inaccurate analysis of trends over time.*

## Data Quality Procedures and Infrastructure

- Initial data validation consists of automated quality checks that ensure data are in the **proper format**.



- **Immediate interim processes for improving data quality** are in place, as larger systemic initiatives are implemented.
- Establish a **data quality office** led by a chief information officer responsible for setting statewide data quality policies, procedures, goals, and technical standards. **Data entry errors are being planned for and addressed by this office.** Data entry errors happen when inaccurate data are entered into a data collection instrument. Data quality office has procedures in place to correct any identified errors.
- State-level personnel play leading roles in **building a culture of data quality among staff at all levels.**
- **Data quality validation is a continuous, inclusive process** that updates all elements of the data system on a regular basis and takes into account both policy and technical considerations.
- **System interoperability.** Data collected in one system are electronically transmittable to other systems. If this criterion is not satisfied, re-entering the same data in multiple systems consumes resources and increases the potential for data entry errors.