

MODULE 6

DATA AND ANALYTICS THE LEARNING NETWORKS GUIDE: BUILDING A LEARNING HEALTHCARE SYSTEM NETWORK

Authors: David Purcell, PhD, Stephanie Loos

Contributors: Jill Schlabig Williams

How to cite this material: *Building a Learning Healthcare System Network. Financial Sustainability, Cincinnati, OH: James M. Anderson Center for Health Systems Excellence, 2020.*



TABLE OF CONTENTS

Introduction	4
Key Elements	4
Registry Data System	4
Data/Analytics Deliverables	5
Data Collection Methods	5
Measures and Analytics	5
Developing Measures for Quality Improvement (QI)	5
Operational Definitions	5
Dashboards, Charts, and Reports	5
Registry and Data Collection	6
IRB Protocol	6
Data Coordinating & Management Center	6
Data Flow Diagram	6
Case Report Forms	6
Phase 1: Design & Development	7
Overview	8
Define Overall Data/Measurement Strategy	8
Create Data Coordinating & Management Center (DCMC) Team and Charter	8
Define Data Collection and Data Flow Strategies	8
Identify and Define Measures: Initial Processes	9
Identify Measures Using Mock Dashboards	9
Identify Chart Types and Create Dummy Reports	9
Draft Measures Table	10





TABLE OF CONTENTS

Create Operational Definitions	10
Initial Registry-Related Processes	11
Begin IRB Protocol	11
Create Case Report Forms and Data Collection Instruments	12
Draft Registry CRF Specifications and Measure Programming Specifications	12
Refine and Finalize Measures and the Data/Measurement Strategy	12
Test Data Collection and Measures	12
Collect Interim Data	13
Phase 2: Implementation & Yearly Cycles	14
Registry Goes Live	15
Conduct Regular Analysis of Data	15
Learn from Variation in Your Data	15
Examine Your Family of Measures	15
Check for Data Quality	15
Data System Revisions and Additional Phase 2 Measures	15
Your Improvement Suggestions	16
Resources and References	16



INTRODUCTION

Quality improvement is inherently data-driven work. Data are used to determine if outcomes are improving and if processes are working as intended. Thus, a strong data and analytics strategy is essential for achieving a learning network's goals.

This module provides an overview of the key elements, processes, and personnel necessary for a robust data system to support a learning health system (LHS).

KEY ELEMENTS

This section outlines the elements that form the core of a robust data/analytics system for an LHS.

Registry Data System

Learning health systems often use an electronic registry to collect data. An electronic registry is a collection of digital information about individuals, usually focused around a specific diagnosis or condition. Generally, the data coordinating and management center (DCMC) and improvement advisor (IA) work with a registry vendor to add the desired measures to the registry or identify variables desired among those already collected in the registry. Depending on the resources and capabilities available to the LHS, these measures can often be transmitted directly from the electronic health records (EHR) of a hospital or clinic.

The registry vendor then incorporates desired measures into the registry. The DCMC provides the vendor with a template data format that should include a description of the structure and content of datasets the vendor will transmit to the data coordinator (DC). At a regularly scheduled data cut-off time each month, the vendor transmits the data to the DCMC.

Desirable technical design features of a registry include:

- Collect data once
- Aggregate measures at various levels (e.g., site, network, project, state)
- Displays data in appropriate QI format (e.g., run charts, small multiples, Pareto)
- Makes data available in timely fashion (real-time data are essential)
- Feedback is sustained
- Easy to report to leaders.

Data/Analytics Deliverables

Data analysis deliverables for an LHS include tracking reports, descriptive summaries, run charts and control charts, statistical analyses, or any other table, listing, or figure that contains data summarized in some fashion and which was derived from a data collection instrument. The analysis specifications for these deliverables should be written down in simple terms or demonstrated using previous examples.

It is primarily the responsibility of the IA to provide the analysis specifications for these reports, which should be reviewed with the DCMC staff. It is the responsibility of the DCMC staff to create technical specifications for the programming of these deliverables.



Data Collection Methods

Potential data collection methods include:

- Electronic health records, which may come via vendors such as Epic or Cerner
- Linkages to external data systems, such as REDCap
- Patient Reported Outcomes (PRO) or mobile health (mHealth) devices
- Manual data entry.

Measures and Analytics

Developing Measures for Quality Improvement (QI)

Many people come to quality improvement (QI) work from more traditional research projects. It is important to recognize that QI measures differ from research measures in key ways. QI data collection should be:

- Driven by the Model for Improvement and the project's key driver diagram (KDD)
- Focused on the SMART (specific, measurable, actionable, relevant, time bound) aim (e.g., "To reduce patient wait time from the 2017 baseline by 25% by January 1, 2019")
- More narrowly focused (e.g., collect ten key measures associated with your KDD, not 100)
- As close to real-time as possible in order to detect change.

QI measures include:

Outcomes measures, which are the voice of the patient, or what the patient experiences. They tell us how the system is performing.

Process measures, which describe the workings of the processes that are logically linked to obtaining the outcomes. These measures address how key parts or steps of the system are performing. Process measures are usually proximal in determining the cause and effect of interventions.

Balancing measures, which provide insight into what happened to the system as we improved the outcomes and processes.

See Module 4: Quality Improvement for more information on and examples of QI measures.

Operational Definitions

Each measure should have a corresponding operational definition (often informally referred to as an "op def") that delineates the key qualities of the measure. A helpful operational definition will include a description of the measure, a definition of the population, the data sources involved, the sampling and data collection plan, the exact calculation used to create the measure, and any relevant information or notes. See Appendix 6.1 for an example.

Dashboards, Charts, and Reports

Visualizing data is essential for quality improvement work at both the network and site levels. Early on, the team should plan to identify which of the following will help drive improvement:

- Dashboards that illustrate the family of measures at a glance
- Charts for key measures (run charts or control charts)
- Reports for other measures or indicators such as missing data or quarterly summaries



Registry and Data Collection

IRB Protocol

The Office for Human Research Protections (OHRP) in the U.S. Department of Health and Human Services has a [helpful guide on IRB](#).

Data Coordinating & Management Center

The data coordinating & management center (DCMC) is responsible for driving a network's approach to data collection, measurement, and management of the data. The DCMC will establish a committee of individuals committed to the network's data and measurement operations. The DCMC will also assume the responsibility for overseeing the data governance process of the network and ensuring that the governance process follows the strategic focus of the network as it relates to quality improvement, research initiatives and sharing. See Module 1: Systems of Leadership for more guidance on data governance.

Data Flow Diagram

The data flow diagram is a graphical depiction of all of the data sources showing the data collection method to be used as well as where the data goes. This diagram should include details regarding where data is stored, how it is integrated (if there are multiple data sources), and where the data analysis deliverables exist.

Case Report Forms

A case report form (CRF) is a printed or electronic document that is designed to record all relevant clinical information on each participant in a QI or research project. The CRF facilitates the collection of standardized data across study participants.





PHASE 1:

DESIGN & DEVELOPMENT



Overview

The registry/measurement process flow diagram (Appendix 6.2) serves as a roadmap for the design and development of an LHS data and measurement system. The roadmap involves many phases and activities among team members. It is important to note that the process is generally iterative, that is, as measures are tested and revised, many of the associated parts (e.g., operational definitions, dashboards, and case report forms) will be revised as well. As such, the timeline and milestones shown here are meant to provide a high-level overview of the key steps, and do not represent a perfectly linear process.

The process as mapped is designed around the beginning of the registry build at time zero (see Appendix 6.2, step #14). The preceding steps are all necessary for the registry build to begin. The registry/measurement process flow diagram provides an opportunity to plan out how many weeks before the beginning of the build each step should take place.

Define Overall Data/Measurement Strategy

Create Data Coordinating & Management Center (DCMC) Team and Charter

After design day, the DCMC team evaluates the projected resources needed with the DCMC director. As soon as practical after the project kickoff, data subgroup team meetings are established in coordination with the improvement advisor (IA), project manager (PM), and project director (PD). Ideally, the regular meetings for the overall project team will include a monthly time slot for DCMC team members to participate, provide updates, and receive information pertaining to overall activities of the project.

The time between the project kickoff meeting and the first learning session (LS1) is the critical period for the work of the DCMC, as the majority of DCMC tasks for a project occur in this time period. These tasks include but are not limited to: developing the necessary data collection forms; implementing and testing the data collection system; and identifying and producing the data analysis deliverables needed for LS1. At the time of the kickoff meeting, the status of the data collection instruments ranges from being undefined to having a draft or previously used instrument to modify or incorporate directly into the current collaborative. Thus, it is critical to coordinate the scheduling of regular data subgroup team meetings so that all planning and implementation activities for reporting can be identified in a timely manner. Of particular importance is to completely specify the scope of work to ensure that the project needs will be met.

See Appendices 6.3 for a DCMC charter template, and 6.4 for an accompanying checklist.

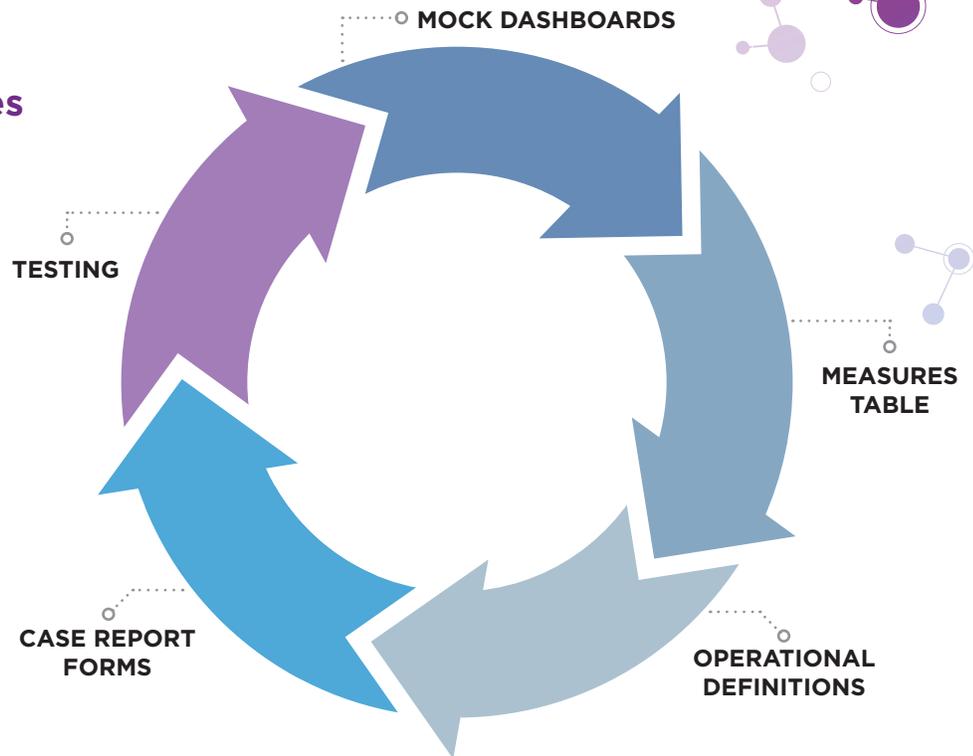
Define Data Collection and Data Flow Strategies

The data collection strategy depends on how many times the data collection instrument is used, who will collect the data, and how complicated or extensive the data collection instrument is. Other choices for the data collection strategy include external data entry systems, electronic health record (EHR) systems, or manual paper forms.



Identify and Define Measures: Initial Processes

This section highlights the key steps involved in identifying and defining measures (see Appendix 6.2, steps #3-6). At a high level, it is helpful to think of the measures development process in this way:



Identify Measures Using Mock Dashboards

Early in the design phase, teams will often have a large number of measures under consideration, far more than is typically needed for a quality improvement (QI) project. This point is especially true for those coming from traditional research where hundreds of variables may be collected for a study. Creating mock dashboards of measures early in the design phase helps the team identify the most essential measures and visualize results for the family of measures. Dashboards should be created at both the individual site and overall network levels. At this point, it is acceptable to simply sketch out a dashboard with pencil and paper or to mockup simple charts in PowerPoint. Ideally, each dashboard will ultimately include no more than 15 measures. See Appendix 6.5 for an example of a mock dashboard for early in the network's development.

Two additional concepts are important to consider at this stage of measure development:

- **Priority Measures:** Which measures are the most important to collect immediately? The KDD should provide guidance, directing the team to the outcomes identified in the Smart AIM, as well as key process measures. We recommend that new projects begin with no more than 5-7 priority measures.
- **Data Burden:** Clinicians and clinical sites are often overwhelmed with the amount of data they collect for various purposes. Teams should keep this in mind as they identify measures for collection and consider questions such as:
 - How difficult will it be for clinicians to collect the data (or for a patient/family to submit it)?
 - Are there existing measures that can be used or repurposed?
 - How often can the measure be collected?
 - Are the questions associated with the measure likely to have a high rate of completion/response?

Identify Chart Types and Create Dummy Reports

As the measures table is refined, the team should define the appropriate chart types for each measure and create dummy reports. Similar to the mock dashboards, dummy reports often help teams think through the idiosyncrasies of measures and operational definitions. Excel is helpful for this step. Teams should review the mock charts and reports with clinical and QI leaders.



Draft Measures Table

The key driver diagram (KDD) and the mock dashboards provide a map for the project's family of measures. A measures table enables the team to see the details for each measure at a glance, including:

- Measure number and ID
- Title
- Description
- Type (outcomes, process, or balancing)
- Data source
- Numerator
- Denominator
- Calculation
- Release order (e.g., Phase 1 or Phase 2)
- Chart type (e.g., U-chart)
- Direction for improvement in measure (e.g., improvement = upwards on the chart)
- X-axis display
- Y-axis display
- Y-axis range
- Inclusions/exclusions

See Appendix 6.6 for a measures table template.

Create Operational Definitions

Following the mock dashboard and measures table, the next step is to create operational definitions for each measure. An operational definition outlines the key qualities of the measure and should include the following:

Description: What does the measure capture? What will it be used for?

Population definition: What patient or family populations will be included in the measure? This should include inclusion and exclusion criteria: e.g., this measure includes only patients under 18 years of age.

Data sources: Where will the data come from?

Sampling and data collection plan: Will sampling be involved? How will the data be collected by the clinician or the site?

Calculation: This element clearly defines the numerator, denominator, and how the measure will be calculated.

Resources: Research citations and key references should be included.

Notes: Any additional relevant notes that will be helpful for understanding the measure should also be included.

See Appendix 6.1 for an example.



Initial Registry-Related Processes

In parallel with the identification and definition of measures, charts, and reports, teams should begin the initial registry-related processes (see Appendix 6.2, steps #3-7). This section highlights the key steps. Appendices 6.7 (an example of a Registry Implementation Plan) and 6.8 (LHS Registry Roles and Responsibilities matrix) are helpful tools at this phase in the network's development.

Begin IRB Protocol

Most learning networks use data for both research and QI purpose. As such, they need an IRB to articulate the dual use and regulate the use of data for research. For details about IRB protocol and template, see Module 2: Governance and Management Module.

When the LHS is determining if they want to pursue an activity that could be QI or research, it is important to review the differences between QI and research. The following table lists the guidelines used within the Anderson Center at Cincinnati Children's Hospital Medical Center to decide when approval is needed from the Institutional Review Board for quality improvement activities:

	Research	QI
Purpose	Test a formal hypothesis.	Assess a process, program, or system.
Starting Point	A prospectively designed, formal, written research hypothesis. Testing of issues that are beyond current science and unusual experience, such as new treatments.	An established set of standards.
Funding	May have federal or external funding (may have a commercial interest in the use of the results)	Usually limited to internal funding
Study Personnel	The involvement in key project roles of researchers who have no ongoing commitment to improvement of the local care situation (and who may have conflicting interests with the patients involved)	Primarily involves only those individuals directly involved with or responsible for the process/procedure/practice being evaluated
Benefits	Knowledge sought may not benefit subjects involved in study.	Knowledge sought directly benefits process/program/system. The majority of patients exposed to the intervention expect to benefit from the knowledge gained.
Risks/Burdens	May put subjects at increased risk beyond the standard practice.	No increased risk or burden for participants, with exception of possible privacy/confidentiality concerns.
Data Collection	Systematic data collection. Randomization of participants into different intervention groups to control for non-random selection bias.	Systematic data collection. Randomization done to achieve equitable allocation of a scarce resource.
End Point	Answer research question.	Improve the program/process/system.
Testing/Analysis	Determine validity of hypothesis.	Compare the program/process/system to established set of standards.
Monitoring the Implementation of Changes & Feedback to Participants	Delayed or ineffective feedback of data from monitoring the implementation of changes, especially if feedback is delayed or altered in order to avoid biasing the interpretation of data. All the data is collected and then analyzed before the benefits of an intervention or project are determined.	Feedback of data. Ongoing, reiterative process so the procedures are constantly being improved during the implementation.
Intended Result	Share findings with individuals associated with the investigation and individuals not associated with the investigation.	Share findings with only those individuals associated with the process/program/system. Information learned has immediate benefit for the program and/or clients receiving the program.



When is Institutional Review Board approval needed for Quality Improvement activities?

IRB approval may be required when the activity involves some of the following characteristics:

- Seeks to develop new knowledge or validate new treatments rather than to assess the implementation of existing knowledge
- When the methodology employs a standard research design, such as randomization
- When the protocol is fixed with a rigid goal, methodology, population, time period, etc.
- When the funding for the activity comes from outside organizations such as the National Institutes of Health (NIH) or those with a commercial interest in the results
- When there will be a delay in the implementation of results
- When the risks from the intervention to participants are greater than minimal.

From: <https://irb.research.chop.edu/quality-improvement-vs-research>

Create Case Report Forms and Data Collection Instruments

Once operational definitions are completed, the team should begin developing the Case Report Forms (CRF) and other instruments that will be used to collect data. Draft CRFs help ensure that the necessary data collection elements are present to calculate the measures. Typically, CRFs are developed and tested with paper forms. See Appendices 6.9 and 6.10 for examples.

Draft Registry CRF Specifications and Measure Programming Specifications

As the CRFs are being developed, the team should also begin to draft the specifications that will map the CRF fields to the registry fields, as well as the programming specifications for the measures. The registry development team needs to know, for example, that the measure “Number of patients in remission” is created by calculating particular values in certain fields.

Refine and Finalize Measures and the Data/M Measurement Strategy

The next steps continue the iterative cycle. As measures are refined through discussions between the clinical leaders, Quality Improvement Specialists (QISs), patient/family representatives, and others, the associated products—measures table, dashboards, charts/reports, measure calculations, programming specifications, and case report forms—must also be updated.

Test Data Collection and Measures

Once the team agrees that the proposed measures and associated data collection tools are close to being finalized, it is essential to pilot test them in clinical settings. Our experience shows that even a small pilot test—e.g., having two clinicians collect priority measures from several patients each—will help refine the measures and the data collection instruments.

We strongly recommend testing the collection of initial priority measures using paper CRFs. This method requires little investment or preparation time, as clinicians can use paper and pen to record patient conditions and responses. Testing in this manner provides insight into how CRF questions can be refined, how and when data should be collected, and if the measures should be further refined (e.g., testing may reveal a patient exclusion criterion that was not considered earlier).



Collect Interim Data

In the event that there is a gap between when the measures are finalized and when the registry will be ready for data entry, teams are encouraged to collect interim data. A best practice is to have sites collect data on a few priority measures and submit aggregate totals for the numerators and denominators to the DCMC. We've seen the benefits of interim data collection with our mature networks. Such interim data collection:

- Provides additional testing of CRFs
- Refines data collection processes at the site level
- Builds data collection into the site's regular work routines
- Facilitates learning from early data.

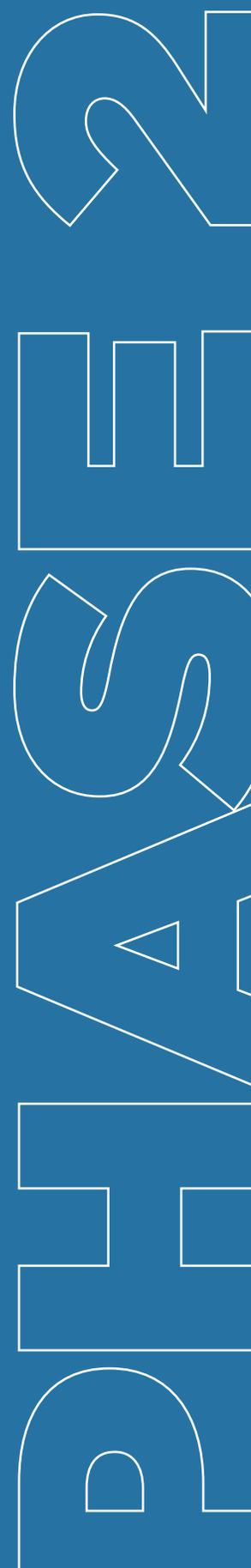
Interim data collection should be kept simple. Existing teams have been successful collecting data with paper forms and simple Excel spreadsheets. Consider testing the interim collection approach with a couple of sites, and then spreading to the entire network. See Appendix 6.7 for an example.





PHASE 2:

IMPLEMENTATION & YEARLY CYCLES



Registry Goes Live

Once a registry goes live, the initial focus should be on ensuring that sites are consistently entering data. Data entry with new network sites can be a challenge as sites adjust to making this process a part of the regular work routines, in addition to data that they may collect for other purposes. Providing coaching and support early on can help sites build momentum and boost their data collection efforts.

Conduct Regular Analysis of Data

System- and site-level dashboards and charts should be reviewed by the DCMC and QIS on at least a monthly basis. Key activities that should be considered are as follows:

Learn from Variation in Your Data

Variation in data is especially important in terms of understanding common cause and special cause variation. In brief, common cause variation is the natural or expected variation that occurs in every process. Special cause variation is the unexpected variation that results from unusual occurrences; you may be able to understand it as being due to a controllable factor.

Project personnel should be aware of the five main rules that determine special cause variation:

1. A single point outside of the upper or lower control limits
2. A run of eight or more points above or below the centerline (note that points right on the centerline do not interrupt a run).
3. Six consecutive points increasing up or down
4. Two out of three consecutive points near a control limit (i.e., between two and three sigma from the centerline).
5. 15 consecutive points close to the centerline (i.e., between the centerline and one sigma)

Examine Your Family of Measures

Consider how your measures are performing in relationship with each other. For example, are your process measures moving in the desired direction and, if so, are you seeing related improvement in your outcome measures? Are your balancing measures—metrics you track to ensure a change to one part of a system is not negatively impacting another part of the system—moving in a worrisome direction?

Check for Data Quality

Check your data at the network level for potential data quality issues such as missing data (e.g., has a particular site not submitted data for some time?) and outliers (e.g., a site may have entered an unusually high number for a patient, and that is now skewing the overall results). In addition, encourage project personnel at the individual sites to focus on data quality at their level as well. See Appendix 6.12 for an example of reports that can be used to verify data quality.

Data System Revisions and Additional Phase 2 Measures

After the registry has launched and initial data collection efforts are off to a solid start, teams may identify areas for improvement to the data system, such as modifying how a measure is calculated or adding charts. Likewise, the team may consider planning for additional measures. Often, the team may want to add measures that were not initially considered to be of high enough priority to include in the initial system launch.



YOUR IMPROVEMENT SUGGESTIONS

We strive to provide the best guide and resources for you. How did we do?

Your feedback helps us continuously improve. Please share your feedback with us: <https://www.surveymonkey.com/r/ZHGJF88>. Thank you!

RESOURCES AND REFERENCES

The following resources are helpful for a deeper understanding of a robust data/analytics system for an LHS:

- Provost, Lloyd P. and Murray, Sandra. The Health Care Data Guide: Learning from Data for Improvement. Josey-Bass, 2011.
- James, Brent. Information System Concepts for Quality Measurement. Medical Care, vol. 41, no. 1, 2003, pp. 171-179. JSTOR, www.jstor.org/stable/3767730.
- AHRQ. Registries for Evaluating Patient Outcomes: A User's Guide: 3rd Edition Addendum. Agency for Healthcare Research and Quality, 2018. <https://effectivehealthcare.ahrq.gov/topics/registries-guide-3rd-edition-addendum/overview>

