

dataCHATT 201:

Introduction to Data Flow and Data Quality Assessment

Mira Levinson, JSI Research & Training Institute, Inc.
Kim Lawton, Quality and Information Management
Lisa Hirschhorn, JSI Research & Training Institute, Inc.

1



Data Quality

- Having quality data is critical for many program activities
 - Clinical care
 - Quality improvement
 - Planning
 - Reporting
- But what do you mean by “quality”, how do you measure it, and why should you care?

2



Presentation Overview

- The Importance of Data Quality for Ryan White Program Grantees
- Essential Steps of Data Flow from Collection to Reporting and Use
- Key Factors for Ensuring Systemic Data Quality
- Key Elements of Data Quality
- Quality Improvement Techniques to Improve Data Quality from Collection through Reporting (a really quick tour)
- Provide an Overview of HAB-Funded Sources of Available TA to Support Data Quality
- Get Participant Feedback

3



The Importance of Data Quality for Ryan White Program Grantees

4



The Importance of Data Quality for Grantees: Data Reporting

- Grantees need to accurately report HIV services provided and patients served to HRSA/HAB
- HRSA needs to accurately reports to Congress for ongoing support of the Ryan White Program

5



The Importance of Data Quality for Grantees: Program Management

- Internal monitoring and evaluation
- Planning
- Quality improvement
- Grant writing

6



Data Quality Concerns

But...

- What if it's not timely?
- What if it's not valid?
- What if it's not complete?
- Why is good data so important to grantees?

7



So where do you start?

- To ensure quality data you need to follow a series of steps in the collection, reporting and use of your data
- These form a flow from identifying what you need to collect through where you will get it to how you will collect and report your data

8



Essential Steps of Data Flow from Collection to Reporting and Use

9



Data Flow Steps: An Overview

- | | |
|---|---|
| 1. Identifying and Defining Data Elements:
<i>What do you want/need to collect?</i> | 5. Data Reporting:
<i>How do you submit the data you have?</i> |
| 2. Data Sources:
<i>Where can you find what you need to collect?</i> | 6. Communicating about Data:
<i>How do you use the data you have to inform our program about how you are doing?</i> |
| 3. Data Collection:
<i>How can you get the data you need to collect?</i> | 7. Using the Data:
<i>How do you use the data you have to inform our program decisions?</i> |
| 4. Data Validation and Data Quality Procedures:
<i>How do you know the data you get is good and accurately reflects what you are trying to measure or report?</i> | 8. Assessing the Effectiveness of the Current System
<i>How can you improve our data system in order to effectively accomplish steps 1 – 7?</i> |

10



Focus on Data Validation and Data Quality Procedures

Efforts to measure and improve data need to happen during all of these steps.

This presentation focuses on Step 4: Data Validation and Data Quality Procedures

11



Key Factors for Ensuring Systemic Data Quality

12



Review and use your data

- Know your data - The best way to improve data quality is to review and use the data!
- Create a system for data quality assessment that is routine, comprehensive and reflective
- Define and follow your data flow steps to collect and report the data

13



Involve your staff

- Engage your staff and your contracted providers in the efforts to ensure data quality!
- Define roles and responsibilities at all levels
 - Consider identifying one or more individuals to oversee data quality procedures (reviewing definitions, protocol development, training, etc).
- Conduct routine training to review data-related procedures and learn about any changes

14



Develop and communicate your requirements and expectations

- Provide routine training to internal staff and contracted providers on reporting requirements, timelines and expectations (through policies, procedures, contracts or MOUs)
- Provide written guidance, and make sure everyone has access to it

15



Ensuring Consistency

- Standardize forms/tools across data collection and reporting efforts
- Develop a written protocol (you own user guide) to document which explains your procedures for data collection, quality and reporting
 - Includes clear and consistent definitions of the key elements for data collection
 - Provides the details for each variable (data source, how you will collect it)
 - Defines who will be responsible for what
 - Is clear and easy to understand
- Develop data review and data cleaning procedures to be performed at all levels
- Update tools and protocols regularly

16



Key Elements of Data Quality

17



Elements of Data Quality

- Validity
- Reliability
- Completeness
- Timeliness
- Integrity
- Confidentiality

18



Validity

Valid data are accurate data defined as “They measure what they are intended to measure.”

19



Validity Questions: Data Collection

- Does the setting and how the questions are being asked potentially compromise their validity?
 - For example: asking an adolescent about sexual activity in front of their parent

20



Validity Questions: Data Collection

- How is the primary data collection and entry being done? Is there potential for error? For example:
 - Client fills out a paper form and misunderstands a question
 - Administrative staff enters form into EMR, and makes an entry error based on client handwriting
 - Databases are not linked, so data must be extracted and then entered hand into HIV program's database: opening opportunity for mistakes.

21



Validity Questions: Data Reporting

- If you are combining data or calculating rates...
 - Are the correct formula and approaches being applied?
 - Are they applied consistently (e.g., from site to site, over time)?
- Are final numbers reported accurately (e.g., does the total add up)?

22



Validity: Steps to Limit Errors

- **Training:**
 - Are all staff trained on definitions and how to complete data entry fields?
- **Validation Checks:**
 - Do the data fall within acceptable range?
 - Look for outliers
 - e.g. age >100
 - CD4 count > 4,000
 - pregnant men

23



More Steps to Limit Errors

- **Validation Rules:**
 - Do you have data validation rules (e.g. can not enter pregnancy if client is male)
- **Validation Activities:**
 - You can do chart extraction to validate data entered
 - Double entry usually reserved for research or when data quality is a significant concern or new staff

24



Example: Validation Checks

In this example Specimen Source:
cervix/endocervix is checked against Gender:
Male

Specimen Source is Cervix, which should only be selected for Female patients.

DATAREG	LAB_ID	CLIENTNO	FACILITY_ID	SPECIMEN_ID	SEX	Gender	Specimen_Source Text
	08st7996		9371	1	Male	Male	Cervix/Endocervix
	08st7996		9371	1	Male	Male	Cervix/Endocervix
	08st7997		9371	1	Male	Male	Cervix/Endocervix
	08st7996		9371	1	Male	Male	Cervix/Endocervix
	08st7995		9371	1	Male	Male	Cervix/Endocervix

25



Reliability

Reliable data are measured and collected consistently (i.e., repeated measurements using the same procedures get the same results)

26



Reliability: Key Questions

- Where are there potential gaps in the data flow which may compromise reliability?
 - The same instrument is no used year to year or across sites
 - Data collected changes without true change in services
 - One site uses a nurse to extract from a medical record, while another uses an non-clinically trained data clerk

27



Procedures to Ensure Reliability

Are steps being taken to limit reliability errors?

- Training
 - Do you provide clear and consistent training across all sites?
 - Is the instrument always administered by trained staff?
- Guidance/Instructions
 - Do you provide detailed procedures and instructions to all sites and providers?
 - Are all providers trained to ask clients to self-identify their ethnicity, race and gender? Is it possible that some providers make assumptions based on appearance?
- Consistent tool (across all sites and providers)
- Refer to user manual

28



Completeness

Complete data do not have any missing elements and are collected on the entire population outlines in the user manual or guidance.

29



Completeness: Key Questions

- Percent of all fields on data collection form filled in
- Percent of all expected reports actually received
- Are the data from all sites that are to report included in aggregate data? If not, which sites are missing?
- Is there a pattern to the sites that were not included in the aggregation of data?

30



Procedures to Ensure Completeness

- Develop a procedure to routinely look for frequency of missing data elements
 - Check for completeness and communicate edits on a routine basis (e.g. monthly)
- Develop and implement procedures follow-up on missing data
 - Volume of missing data often diminishes over time once staff are aware that someone is looking at it
 - Procedures may be different for data received from contractors versus internally collected data
 - electronic data submission vs. paper data submission

31



Look for “missing data” trends

- Look for trends in missing data, and ask “why?”
- Are there barriers to capturing or entering the data?
 - Meet with your staff and ask for their insights
 - Use this information for data collection planning

32



Timeliness

Timely data are...

- sufficiently current and frequent to inform management decision-making
- received by the established deadline
- received with adequate time to review for other elements of quality, and to address identified gaps

33



Timeliness: Key Questions

- Is a regularized schedule of data collection in place to meet program management needs? When are your established deadlines?
- Does program staff and contractors know and understand the reporting deadline? Is it consistent across all reporting sites?

34



Timeliness: More Key Questions

- Is there adequate time to review data for other aspects of quality and address identified gaps before it is needed for reporting or other use?
- Are data available on a frequent enough basis to inform program management decisions?
- Are data being collected and reported according to your timeline?

35



Optimal Timeline for Collecting Data to Ensure Quality

- Work back from the submission deadline
 - include time to review, address identified gaps, etc.
- More frequent collection allows for more time to review data collected
 - Care and services being provided
 - Missing data
 - Other data problems
- Grantees with subcontractors can request data submissions more frequently than reporting requires (more than annually)

36



Procedures to Ensure Timeliness

- Define and set reasonable timelines
- Communicate and stick with timelines
- Include a process for reviewing whether data was submitted on time, providing feedback and requesting revisions
- Consider implementing consequences for lateness

37



Integrity

Data are protected from deliberate bias or manipulation for any reasons

38



Integrity: Key Questions

- Are there risks that data might be manipulated for any reasons?
- What systems are in place to minimize such risks?

39



Confidentiality

Clients are assured that their data will be maintained according to organization, state and national standards

40



Confidentiality: Key Questions

Do you provide routine training...

- to program staff on the importance of confidentiality, and on confidentiality requirements and procedures?
- to IT staff on the specific issues of HIV confidentiality and electronic information storage and transfer?
- to contracted service providers on procedures for data submission?
- to clients on confidentiality procedures?

41



Procedures to Ensure Integrity and Confidentiality

- Training
 - Train all staff and contracted providers on confidentiality and privacy protocols
- Electronic Data Security
 - Document user access to database
 - Limit user access to database
 - Consider security limitations of laptops, handheld devices, etc

42



Procedures to Ensure Integrity and Confidentiality

- Security of Paper Data
 - Store paperwork in a secure, locked cabinet and/or user-restricted area
- Inform Clients of Confidentiality and Privacy Protocols

43



Discussion: How Does This Apply To Me?

- Validity (accuracy)
- Reliability (consistency)
- Completeness (all there)
- Timeliness (there when you need it)
- Integrity (honesty)
- Confidential

44



Ensure Quality Through Assessment

- When to assess program data quality:
- Integrate data quality control mechanisms into standard operating procedures and software
 - Integrate data quality checks into routine supervisory or contract monitoring visits
 - Conduct periodic formal assessments
 - Provide feedback on submitted data

45



Quality Improvement Techniques to Improve Data Quality from Collection through Reporting (a really quick tour)

46



Application of Basic Data Quality Improvement (QI) Techniques

- The same concepts apply to improving data quality as they do to improving quality of care:
 - Measure the quality.
 - Explore steps required for quality data and where gaps may have occurred (flow chart).
 - Understand the potential causes of the identified gap (fishbone or cause and effect).

47



A Sample QI Technique: Plan-Do-Study-Act Cycle

Plan: Develop a QI Project goal (i.e. what you want to accomplish) based on assessment of data quality

- Decrease missing data, improve timeliness,
- Form a team
- Identify where you think the problem (gap) may be and develop a potential solution

Do: Carry out the proposed solution

Study: Analyze your data, summarize what was learned, compare with what you wanted to achieve-did the solution work

Act: Determine next steps (if worked, how to adapt from the American Heart Association)



48



Example: Low Reported Pap Smear Rates

PLAN

- Identify the problem: A hospital-based site notices that their Pap rates for HIVQUAL are 75%, but those reported in the RDR are only 40%.
- Develop a QI Project Goal: They want to improve the quality of reported data.
- Form a Team: A team is formed including the program data manager, a nurse provider, and a case manager.
- They define the goal as decreasing the difference between reported and actual rates to less than 10%.

49



Example: Low Reported Pap Smear Rates

PLAN

- Identify the data steps required for a Pap smear to be included in the RDR report
 - Internally: internal lab results are automatically entered into the EMR, which is then used to download data into a program database for RDR submission
 - Versus HIVQUAL: chart review of client sample and entry into HIVQUAL database

50



Example: Low Reported Pap Smear Rates

PLAN

- Pap results for patients seen by external providers are not received 25% of the time.
- When these results are received, they are manually entered into a different field than the one used for Pap results for patients seen by internal providers (done via automatic transfer from lab system).
- For HIVQUAL reviews, both fields are manually extracted, but the automated RDR report only extracts the data field of the program database of the internally-provided Pap tests.

51



More Plan, Do, and Study

- **PLAN:** Modify IT systems can be modified so that data sources are the same OR reporting draws from both Pap data sources.
- **DO:** Ask the hospital IT department to reprogram so that external Paps can go into the same field OR the RDR report can look at both fields
- **STUDY:** nothing happens as the hospital EMR is a proprietary software and takes significant resources to revise and will take many months

52



Act and the Next Cycle

- **ACT:** Decide to try a different approach for an interim solution
- **PLAN:** Establish a Log for women getting Paps from external providers and use to manually enter into program database.
- **DO:** train a nurse and data manager to use an Excel spreadsheet to enter any woman getting a Pap from provider external to the clinic and educate all providers to give the Pap results to the nurse before sending to medical records
- **STUDY:** Next RDR rate is only 18% below HIVQUAL data.
- **ACT:** Continue log and also work with PO to get resources to ultimately automate capture of externally provided Pap tests.

53



Provide an Overview of HAB-Funded Sources of Available TA to Support Data Quality

54



TA Resources to Support Data Quality

- Project officer
- TARGET Center
 - <http://www.careacttarget.org>
- dataCHATT
 - <http://www.datachatt.jsi.com/>
- Ryan White HIV/AIDS Program Data Report TA
 - <http://datasupport.hab.hrsa.gov/>
- CAREWare TA
 - <http://hab.hrsa.gov/careware/>
- National Alliance of State and Territorial AIDS Directors Cooperative Agreement (NASTAD)
 - <http://www.nastad.org/Programs/hivcareandtreatment>
- National Quality Center (NQC)
 - <http://www.nationalqualitycenter.org/>
- Manuals (for example, the HIPAA Manual)
 - <http://hab.hrsa.gov/publications/hippa04.htm>

55



Get Participant Feedback

56



Data Academy

- dataCHATT is developing a series of web-based training modules.
- This Data Academy will include training modules on data collection, data quality, data reporting and using data.
- We need your feedback to make sure the information is presented effectively.

57



Feedback

- Was this content useful?
- Appropriate?
- Did it meet your needs?
- Any suggestions?
- Can we contact you to review future Data Academy modules?

58



Acknowledgements

- The HIV/AIDS Bureau for support of this Cooperative Agreement
- The 101 Grantees who participated in the Request for Information
- JSI contributing staff (Julie Hook, Kim Watson, Michael Rodriguez and the dataCHATT team)
- Positive Outcomes, Inc.

59



For more information...

Visit the dataCHATT website:
www.datachatt.jsi.com

For copies of today's presentation, contact us at: datachatt@jsi.com

60


