

ACT Health Data Quality Framework

Version: 1.0
Date : 18 December 2013

Table of Contents

Table of Contents	2
Acknowledgements	3
Document Control	3
Document Endorsement	3
Glossary	4
1 Introduction	5
1.1 Data and Quality Improvement within a Health System	5
1.2 The Information Management Framework	6
1.3 What is the Data Quality Framework?	7
1.4 Purpose and Scope	7
1.5 The Impact of Data Quality on ACT Health	8
1.6 What is Data Quality?	9
1.7 Aboriginal and Torres Strait Islander Data	9
2 The Data Quality Framework	10
2.1 Data Integrity Strategy	10
2.2 Data Quality Policy	10
2.3 The Data Quality Improvement Cycle	10
2.3.1 Five stages of the Data Quality Improvement Cycle	10
2.3.2 Roles & Responsibilities	11
2.4 Data Quality Assessment Tool	13
2.4.1 Dimensions of Data Quality	15
Institutional Environment	15
Relevance	15
Timeliness	16
Accuracy	17
Coherence	18
Interpretability	18
Accessibility	19
2.4.2 Applying the Data Quality Assessment Tool	19
2.4.3 ACT Health Data Quality Indicators	20
Data Quality Indicator	21
3 References	23
4 Appendices	24
{ Appendix A - Data Integrity Strategy }	24
{ Appendix B – Data Quality Policy }	24
Appendix C – ACT Health Data Quality Statement	25

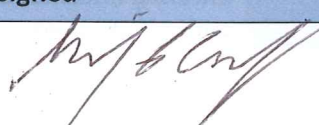
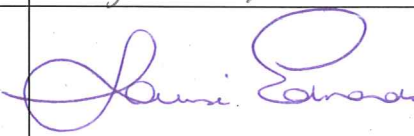
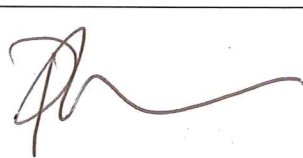
Acknowledgements

This Data Quality Framework strongly draws from the Australian Bureau of Statistics (ABS) and Canadian Institute of Health Informatics (CIHI), respective data quality frameworks. The Data Governance and Standards Unit, Performance and Information Branch, ACT Health would like to recognise and thank the ABS, CIHI, and Australian Institute of Health & Welfare (AIHW) for their input in development of this framework. Additionally, the input and feedback of ACT Health staff in development and testing of the concepts has been significant in providing a more robust and practical outcome.

Document Control

Version	Date	Author	Comments
0.1	12/04/2010	S. Brown	Document Initiated
0.2	04/05/2010	S. Brown	Document Updated
0.3	27/05/2010	S. Brown	Document Updated following review by the Chief Information Officer and an ACTHEIM Working Group workshop
0.4	15/06/2010	S. Brown	Document Updated following further review by ACTHEIM Working Group and the Data Management Committee
0.5	18/06/2010	S. Brown	Data Quality Indicator adjusted following review by CIO
0.6	01/02/2011	S. Brown	Some changes following pilot
0.7	01/02/2013	P. Crawford	Major review following workshop on pilot.
0.8	04/02/2013	L. Edmonds	Editorial changes and inclusion references.
0.9	04/03/2013	P. Crawford L. Edmonds	Review and update following comments reviewed from Data Standards Working Group members
0.91	21/05/2013	P. Crawford	Changed 'Health Directorate' to 'ACT Health'.
0.92	12/08/2013	P. Crawford	Added Section 1.7 Aboriginal and Torres Strait Islander Data, following feedback from the Aboriginal and Torres Strait Islander Health Unit
1.0	06/01/2014	P. Crawford	Endorsed by Executive Council

Document Endorsement

Position	Name	Signed	Date
Senior Information Analyst, Data Governance and Standards Unit	Phil Crawford		12/8/2013
Data Standards Working Group Chair. Senior Manager, Data Governance and Standards	Louise Edmonds		12/8/13
Information Management Committee Chair. Executive Director, Performance and Innovation Branch	Phil Ghirardello		13/8/13
Executive Council	Dr Peggy Brown (Chair)	N/A	18 Dec 2014

Glossary

Term	Definition
ABS	Australian Bureau of Statistics
AIHW	Australian Institute of Health & Welfare
Business Unit	A corporate, clinical, or program area that works within the boundaries of a Portfolio Executive Team within ACT Health.
CIHI	Canadian Institute for Health Informatics
Data	Distinct pieces of information, usually formatted in a special way. Data can exist in a variety of forms - as numbers or text on pieces of paper, as bits and bytes stored in electronic memory, or as facts stored in a person's mind.
Data Collections	Data that has been systematically collected through the process of providing health services to consumers, then collated into meaningful groups and stored by the Directorate.
Data Manager	Also commonly referred to as Data Stewards, the Data Managers are responsible for reporting, manipulation and storage of data. They generally work to a specific area, dataset, or data collection.
Data Modelling	The process of conceptually and graphically documenting a database's entities, the relationship between these entities, and their attributes.
Data Producer	The areas, or business units, where data have been produced for a dataset.
Data Source	Where the data originated from. For this document, the purposeful definition is the information system, database, users, and business unit that the data has originally come from.
Data User	A stakeholder or client of the datasets, or outputs of data collections.
Dataset	A data file or collection of inter related data
Metadata	Information about data. Where the contents of a field within a system are described as data, metadata would be the attributes of that field and data. For instance – a definition on the data and its format, character make up, or input options.
METeOR	AIHW's Metadata Online Registry -the National Health Data Dictionary
User	A user is generally the person who is using computer systems or entering information into an information system, not to be confused with a Data User.

1 Introduction

1.1 Data and Quality Improvement within a Health System

The Australian health system is challenged with increasing cost and demand pressures due to a range of factors but particularly, an ageing population, increasing prevalence of chronic disease and shortages in skilled healthcare workers. This environment requires effective use of resources within that system leveraging every interaction between patients and care providers to achieve maximum impact on health outcomes. The demand for information to measure these outcomes and the effectiveness of the system that delivers these services is significant.

Critically, the data that underpins this service delivery needs to accurately reflect those interactions. Data must be reliable in order to support quality improvement and in understanding health care delivery.

Data quality is a fundamental requirement in health systems. Health personnel require reliable consistent and coherent data to ensure appropriate patient care, planning and management decisions can be made. An integrated organisational approach that supports a culture of data quality and engagement of staff on all levels is required to meet the needs of a data driven organisation.

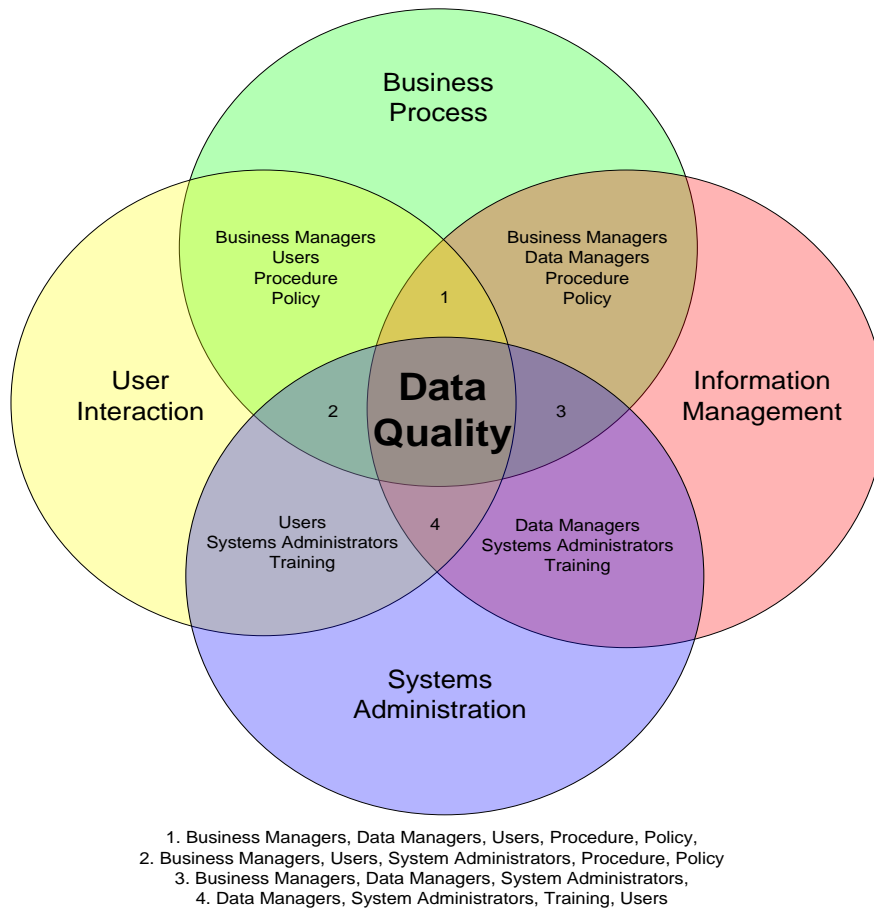


Figure 1.1: The integrated organisational approach required to sustain data quality.

1.2 The Information Management Framework

The *Health Directorate Information Management Framework 2012-2014* articulates the six information management principles that inform the Directorate's approach to managing information.

Information Management Principles

- **Purpose** - Authoritative information is provided to meet the needs and priorities of the ACT Government and the people that use our services.
- **Governance** - Information is managed in an ethical and efficient manner to support patient care, planning, research and service delivery.
- **Collection** - Information is collected and collated efficiently and effectively to minimise the burden on respondents.
- **Quality** - Quality information is produced that is relevant, timely, accurate, coherent, transparent and accessible.
- **Use and Disclosure** - The analysis, interpretation and reporting of health and health-related information occurs in accordance with relevant legislative and privacy codes.
- **Security** - Information is protected and preserved throughout the information lifecycle.

The information management principle of quality determines that:

Quality information is produced that is relevant, timely, accurate, coherent, transparent and accessible.

The principle is further supported by three sub-principles:

- Information is of a quality that makes it fit for purpose.
- Information about data (Metadata) is developed and is accessible to all users of information; and
- Correction of data is done in or as close as possible to the source system.

The Data Quality Framework provides an organisation wide approach to applying the quality principle within the Information Management Framework.

1.3 What is the Data Quality Framework?

The Data Quality Framework provides an objective approach to assessing and improving the quality of data created and managed by ACT Health. The framework encompasses four main components.

Data Quality Framework Components

- The Data Integrity Strategy
- Data Quality Policy
- The Data Quality Improvement Cycle
- Data Quality Assessment tool

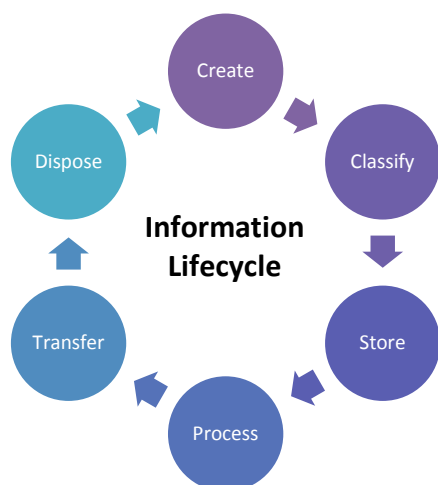
The framework provides quality assurance tools supporting quality process with the input, output, throughput, use, manipulation, or reporting of data. It also provides assessment tools to aid in the analysis of the fitness for purpose a data collection may have to a particular report or usage.

1.4 Purpose and Scope

The framework can be applied to all data collections within the Directorate including databases, systems, registries and reports. The framework applies to all staff members that create, modify, use and interpret data including executive, management, data management, clinical and administration staff.

The framework describes tools to help organise, assess, and incorporate activities related to data quality and provides awareness of the impact various dimensions may have on data quality. Staff members who make decisions based on available information will find value in the provision of robust data using the data quality framework and its tools.

Data quality practices should be maintained through the full lifecycle of data within the Directorate.



It is important that users endeavour to enter data into systems and registries correctly in the first instance. The effort and consequently cost of amending data downstream is significant and potential issues may arise from decisions made on what is subsequently flawed data.

Due to the scope of the framework, it is necessarily general in description and focuses on issues that affect almost all data users, databases, data sources, or information systems within the Directorate.

1.5 The Impact of Data Quality on ACT Health

High quality data underpins high quality information being available for decision making. Data entered into information systems across the Directorate are used in assessing patient needs and care, service delivery to patients/clients, funding, management, planning and strategic decision making. Without reliable consistent data, decisions made may not reflect the situation accurately and consequently could be flawed and potentially dangerous. Data quality has implications for a number of areas including;

Improved Patient Care – Quality patient data supports high level care for the patient. Reductions in duplicate registrations support complete patient records being available for clinicians. Improved patient contact information improves communications to the patient and a higher quality service delivery.

Appropriate Funding Allocation – Quality activity and costing data results in more accurate profile of the cost of service delivery within ACT Health. It provides a clearer understanding of services utilisation and resource consumption so that funding can be appropriately allocated to services. In addition the Directorate will be able to claim appropriate levels of funding from various funding bodies including various government or private providers, and cross border care for interstate residents.

Improved Customer Service – Quality data results in a higher quality of analysis and decision making. Efficiencies identified through analysis can support improvements in customer service. Patterns and correlations identified through analysis can lead to improved clinical decisions for individual patients.

Research – Quality data used for research supports more precise or relevant conclusions or clinical decision making. This may then inform a range of outcomes including improved models of care and appropriate and effective funding priorities for the Directorate.

1.6 What is Data Quality?

Quality in the health care setting may be defined as the ‘extent to which a health care services or product produces a desired outcome’ⁱ. ‘Quality improvement is a system by which better health outcomes are achieved through analysing and improving service delivery processes.’ⁱⁱ Quality information at its fundamental level is relevant, timely, accurate, coherent, transparent and accessible.ⁱⁱⁱ

In the past, the term “data quality” has been generally understood as a concept for accuracy. The new consensus is that data are considered to have sufficient data quality when it is appropriate to use for its intended purpose, or “fit for purpose”. The purpose may include operational use, decision making, mandated reporting, and legislative requirements. In this context, fitness implies both freedom from defects and possession of the desired features and attributes for sustainable use.

Data are often used for secondary purposes in addition to what it was initially intended. Thus, meeting “organisational fit for purpose” is the endeavour of a data quality initiative. As the data is entered into the system by a user, system administrators control the systems, data managers manipulate and report on the data, and managers use the data to form decisions. It is crucial to understand data quality as an organisational issue. Data quality cannot be confined to being an issue placed solely into the hands of Information Technology, Information Management, nor fully the responsibility of the business unit or users.

Where data quality and fit for purpose can seem subjective, raising the awareness of organisational fit for purpose and assessment of datasets for appropriate usage is a cornerstone of the framework. Documentation of data quality is essential for both internal and external stakeholders, and this framework will provide the guide and tools to do so.

1.7 Aboriginal and Torres Strait Islander Data

The ACT Government is committed to closing the gap between the life outcomes and opportunities experienced by the Aboriginal and Torres Strait Islander population and the non-Aboriginal and Torres Strait Islander population.

Improving Aboriginal and Torres Strait Islander identification in key administrative data collections is a key requirement to enable accurate assessment of health outcomes of, and service use by, Aboriginal and Torres Strait Islander people.

Good quality data is important to accurately assess the health and welfare of Australia’s Aboriginal and Torres Strait Islander population. Knowing whether Aboriginal and Torres Strait Islanders have adequate access to services, and the extent to which their health and welfare has changed over time and relative to other Australians, is important for assessing the effectiveness of policies and programs aimed at improving Aboriginal and Torres Strait Islander health and wellbeing. However, it is difficult to assess and fully understand such issues if Aboriginal and Torres Strait Islander status is not accurately identified in key administrative data sets.

2 The Data Quality Framework

Data Quality Framework Components

- The Data Integrity Strategy
- Data Quality Policy
- The Data Quality Improvement Cycle
- Data Quality Assessment tool

2.1 Data Integrity Strategy

The Data Integrity Strategy (Appendix A) articulates the direction for data quality improvement in ACT Health. The Strategy describes and prioritises a number of initiatives that strengthen the Directorates approach to information collection, collation, use and security.

2.2 Data Quality Policy

The Data Quality Policy (Appendix B) is the instrument that authorises the framework approach. It focuses on the roles and responsibilities of staff ensuring that data are accurately recorded, updated and maintained over time to support a credible, rigorous source of information. The Data Quality Policy supports the organisation's position of a data driven decision culture based on a consistent and rigorous approach to data quality.

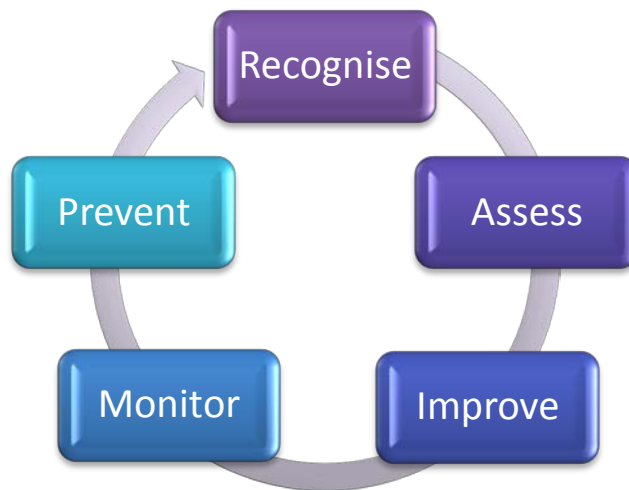
2.3 The Data Quality Improvement Cycle

The data quality improvement cycle addresses the processes, models and methodologies used to create or change data management processes as part of a data quality initiative. It is an iterative approach that consists of a number of phases or stages of analysing and assessing input, throughput, and output while continually moving to improve the quality.

2.3.1 Five stages of the Data Quality Improvement Cycle

The quality improvement cycle for the ACT Health framework is a combination of the PDCA^{iv} (Plan, Do, Check, Adjust) and DMAIC^v (Define, Measure, Analyse, Improve, Control) model's for quality improvement. These models are international standard practices and recognised as exemplars in the approach to improvement in any aspect of quality.

The cycle's approach is to **recognise** the issues to be addressed, **assess** the approach, **implement** improvement, and **monitor** the outcome in order to **prevent** future issues. The diagram below depicts a process that is iterative where the cycle is continuously active to ensure that optimum data quality is achieved and maintained.



Recognise: Actively identifying data issues that need to be addressed through a quality improvement process.

Assess: Prepare the processes required to improve the identified quality issue whilst designing any changes that are needed.

Improve: Implement the planned quality improvements by applying them to resolve the issue.

Monitor: Monitor the outcomes to assess whether the cycle needs to be repeated. Maintain the implemented resolution in order to prevent future issues of the same kind.

Prevent: Ensure the cycle is constantly maintained to recognise arising issues and implement a strategy utilising this process. This will ensure there is a constant improvement process.

2.3.2 Roles & Responsibilities

Data quality is the responsibility of every staff member in respect to their core role within the organisation. The day to day activities of staff across the organisation generate an enormous amount of data and information used to inform patient care, operations, management, planning and research. Data quality is intrinsic to our business and thereby becomes each individual's responsibility.

More specific roles and responsibilities are visible in the data quality framework and in the application of tools like the data quality improvement cycle. The data quality improvement cycle is designed to be flexible enough that it can be applied to any system, process, or general quality improvement initiative. It is important to note that a staff member may have more than one role at a time and that their roles may change for different data quality initiatives.

Executives – Senior Executives and Executive Directors

- Lead and promote the development of a data quality culture throughout the Directorate.

Managers –Staff members with a supervisory role.

- Ensure that business units adhere to the data quality framework and policies,
- Provide the necessary resources to manage quality,
- Encourage that initiatives and projects involve a data quality component from the beginning,
- Provide direction and prioritise quality improvement of data assets.

ICT Personnel – Any staff member that provides ICT support. This includes Shared Services and ACT Health staff.

- Address system related data quality issues with the appropriate stakeholders as they are identified,
- Address system issues in a timely manner and ensure that stakeholders are identified of issues that may impact data quality,
- Identify ways to improve data source information systems using the data quality improvement cycle.

System Administrators - Staff members with a system administration role.

- Address data quality issues with the appropriate stakeholders as they are identified,
- Ensure system documentation and metadata meet the data quality framework and data governance framework criteria,
- Work with the appropriate business units to ensure metadata correctly reflects system fields and data elements.

System Trainers – Any staff member that trains others in the use of a system. This includes Shared Services and ACT Health staff.

- Promote data quality and embed best practice into training,
- Include training materials that support best practice data quality.

Data Managers and Analysts – Any staff member that manages or analysis data.

- Address data quality issues with the appropriate stakeholders as they are identified,
- Promote data quality and embed best practice into work practices,
- Ensure that relevant data standards are applied,
- Ensure that data are appropriately sourced and documented,
- Provide guidance on data quality issues and problem resolution,
- Review and report on data quality,
- Conduct root cause analysis of data issues,
- Identify ways to improve data collections and data assets proactively.

Data Quality Officers – All staff where designated data quality functions are part of their role are considered to be Data Quality Officers.

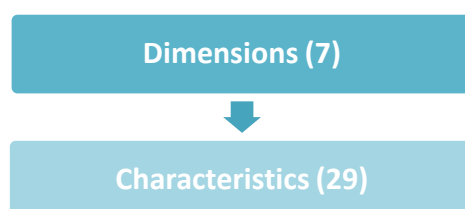
- Assess and provide reports on compliance with the data quality framework,
- Assist in data quality studies or ad hoc data quality reporting or initiatives,
- Assist in development of data quality infrastructure, strategy, and policy,
- Provide feedback to System Trainers to help develop best practice,
- Assist in root cause analysis of data issues.

Clinical and Administrative staff – Any staff member that captures or enters data on a form, in a medical record, or in a system.

- Ensure data are captured and validated at the point of client contact,
- Ensure data are correct, current and complete,
- Ensure data are entered as close to real-time as possible,
- Report data quality issues as they are identified to the appropriate manager,
- Escalate data quality issues that impact on patient safety to the manager as a priority,
- Participated in data quality education activities,
- Address feedback regarding data quality from management and participate in the resolution of issues as appropriate.

2.4 Data Quality Assessment Tool

The data quality assessment tool comprises seven dimensions each of which is divided into a number of data characteristics that may be used to systematically assess the fitness for purpose of data.



ACT Health has adopted the seven dimensions of data quality developed by the ABS^{vi}. All seven dimensions should be used when assessing the quality of data sets, collections and reports in order to support robust decision making and management. Descriptions of each dimension are provided, as defined by the ABS and expanded on by ACT Health.

The seven dimensions are not mutually exclusive, and the quality questions they raise may overlap. The importance of each dimension is also not necessarily equal and may vary depending on the data and its context.

The application of the Data Quality Assessment tool will present different results for different reports and extracts even where they use the same sources system(s). Each report, extract or dataset has its own purpose and scope. In assessing the 'fit for purpose' of each requires a different level of emphasis against each dimension being applied.

The Dimensions of Data Quality

- **Institutional Environment** - The organisational, and business factors, that have significant influence on the effectiveness and credibility of the data being produced. Key aspects of this dimension are; impartiality and objectivity, professional independence, mandate for data collection, adequacy of resources, quality commitment, and statistical confidentiality.
- **Relevance** - The degree by which data meets the real need of clients. Relevance is concerned with whether the available information sheds light on the issues most important to data users. It may be described in terms of key data user needs, key concepts and classifications used, and the scope of the collection (including the reference period). It also reflects the degree to which a data source or collection meets the current and future needs of data users.
- **Timeliness** - The delay between the reference point (or end of the reference period) to which the information pertains and the date on which the information becomes available. This can be described as how current or up to date the data are at the time of release, and can be measured as the gap between reference point and the date the information is available.
- **Accuracy** - Accuracy is what used to be traditionally thought of when the term Data Quality was used, and is the pertinent dimension to data producers. It can be described as the degree by which the information correctly reflects the reality it was designated to measure. It may be characterised in terms of error: traditionally decomposed into bias (systematic error) and variance (random error) components. It is described in terms of sources of error that potentially cause inaccuracy (sampling/response errors, non-response errors, coverage).
- **Coherence** - The degree by which it can be successfully brought together with other data within a broad analytic framework and over time. Coherence encompasses the internal consistency of a data collection.
- **Interpretability** - The availability and clarity of metadata, including concepts, classifications, modelling, data dictionaries, and measures of accuracy necessary to interpret and utilise the data appropriately. This includes the appropriate presentation of data such that it aids in the correct interpretation of the data.
- **Accessibility** - The ease with which data or datasets can be referenced by users. This includes the suitability of the form or medium through which the information can be accessed. This may also include factors such as the cost of producing the data, or the cost of accessing the data.

2.4.1 Dimensions of Data Quality

Institutional Environment

Institutional Environment

- Impartiality and objectivity
- Professional independence
- Mandate for data collection
- Adequacy of resources
- Quality commitment
- Statistical confidentiality

Characteristics

- **Impartiality and objectivity:** whether the production and dissemination of data are undertaken in an objective, professional and transparent manner.
- **Professional independence:** the extent to which the agency producing data is independent from other policy, regulatory or administrative departments and bodies, as well as from private sector operators, and potential conflict of interest.
- **Mandate for data collection:** the extent to which administrative organisations, businesses and households, and the public at large may be compelled by law to allow access to, or to provide data to, the agency producing statistics.
- **Adequacy of resources:** the extent to which the resources available to the agency are sufficient to meet its needs in terms of the production or collection of data.
- **Quality commitment:** the extent to which processes, staff and facilities are in place for ensuring the data produced are commensurate with their quality objectives.
- **Statistical confidentiality:** the extent to which the privacy of data providers (patients, consumers or others), and the confidentiality of the information they provide, are guaranteed (if relevant).

Relevance

Relevance

- Scope and coverage
- Reference period
- Geographical detail
- Main outputs / data items
- Classifications and statistical standards
- Type of estimates available
- Other cautions

Characteristics

- **Scope and coverage:** the purpose or aim for collecting the information, including identification of the target population, discussion of whom the data represent, who is excluded and whether there are any impacts or biases caused by exclusion of particular people, areas or groups.

- **Reference period:** this refers to the period for which the data were collected (e.g., the September-December quarter of the 2008-09 financial year), as well as whether there were any exceptions to the collection period (e.g., delays in receipt of data, changes to field collection processes due to natural disasters).
- **Geographic detail:** information about the level of geographical detail available for the data (e.g., postcode area, Statistical Local Area) and the actual geographic regions for which data are available.
- **Main outputs / data items:** whether the data measures the concepts meant to be measured for its intended uses.
- **Classifications and statistical standards:** the extent to which the classifications and standards used reflect the target concepts to be measured or the population of interest.
- **Type of estimates available:** this refers to the nature of the statistics produced, which could be index numbers, trend estimates, seasonally adjusted data, or original unadjusted data.
- **Other cautions:** information about any other relevant issue or caution that should be exercised in the use of the data.

Timeliness



Characteristics

- **Timing:** this refers to the time lag between the reference period and when the data actually becomes available (including the time lag between the advertised date for release and the actual date of release). For example, the reference period may be the 2004-05 financial year, but data may not become available for analysis until the middle of 2006.
- **Frequency:** this refers to whether the data collection was conducted on a one-off basis, or whether it is expected to be ongoing. If it is expected to be ongoing, frequency also includes information about the proposed frequency of repeated collections and when data will be released for subsequent reference periods.

Accuracy

Accuracy

- Coverage error
- Sample error
- Non-response error
- Response error
- Other sources of errors
- Revisions to data

Characteristics

- **Coverage error:** this occurs when a unit in the sample is incorrectly excluded or included, or is duplicated in the sample. Coverage of the statistical measures could be assessed by comparing the population included for the data collection to the target population.
- **Sample error:** where sampling is used, the impact of sample error can be assessed using information about the total sample size and the size of the sample in key output levels (e.g., number of sample units in a particular geographical area), the sampling error of the key measures, and the extent to which there are changes or deficiencies in the sample which could impact on accuracy.
- **Non-response error:** this refers to incomplete information provided by a respondent (e.g., when some data are missing, or the respondent has not answered all questions or provided all required information). Assessment should be based on non-response rates, or percentages of estimates imputed, and any statistical corrections or adjustment made to the estimates to address the bias from missing data.
- **Response error:** this refers to a type of error caused by respondents intentionally or accidentally providing inaccurate responses, or incomplete responses, during the provision of data. This occurs in administrative data collection where forms, or concepts on forms, are not well understood by respondents. Respondent errors are usually gauged by comparison with alternative sources of data and follow-up procedures.
- **Other sources of errors:** Any other serious accuracy problems with the collection should be considered. These may include errors caused by incorrect processing of data (e.g. erroneous data entry or recognition), alterations made to the data to ensure the confidentiality of the respondents (e.g. by adding "noise" to the data), rounding errors involved during collection, processing or dissemination, and other quality assurance processes.
- **Revisions to data:** the extent to which the data are subject to revision or correction, in light of new information or following rectification of errors in processing or estimation, and the time frame in which revisions are produced.

Coherence

Coherence

- Changes to data items
- Comparison across data items
- Comparison with previous releases
- Comparison with other products available

Characteristics

- **Changes to data items:** to what extent a long time series of particular data items might be available, or whether significant changes have occurred to the way that data are collected.
- **Comparison across data items:** this refers to the capacity to be able to make meaningful comparisons across multiple data items within the same collection. The ability to make comparisons may be affected if there have been significant changes in collection, processing or estimation methodology which might have occurred across multiple items within a collection.
- **Comparison with previous releases:** the extent to which there have been significant changes in collection, processing or estimation methodology in this release compared with previous releases, or any 'real world' events which have impacted on the data since the previous release.
- **Comparison with other products available:** this refers to whether there are any other data sources with which a particular series has been compared, and whether these two sources tell the same story. This aspect may also include identification of any other key data sources with which the data cannot be compared, and the reasons for this, such as differences in scope or definitions.

Interpretability

Interpretability

- Presentation of the information
- Availability of information regarding the data

Characteristics

- **Presentation of the information:** the form of presentation and the use of analytical summaries to help draw out the key message of the data
- **Availability of information regarding the data:** the availability of key material to support correct interpretation, such as concepts, sources and methods; manuals and user guides; and measures of accuracy of data.

Accessibility

Accessibility

- Accessibility to the public
- Data products available

Characteristics

- **Accessibility to the public:** the extent to which the data are publicly available, or the level of access restrictions. Additionally, special data services may include the availability of special or non-standard groupings of data items or outputs, if required.
- **Data products available:** this refers to the specific products available (e.g., publications, spreadsheets), the formats of these products, their cost, and the available data items which they contain.

2.4.2 Applying the Data Quality Assessment Tool

The data quality assessment tool enables a comprehensive and multi-dimensional assessment of the quality of data within the Directorate. It is intended that the tool enables data users, data managers and data providers to:

- assess the quality of a data item or a collection of data items, with reference to the user's specific purpose and requirements; and
- design a data collection or product which is fit for purpose.

Application of the tool by suppliers and users of data

It is recommended that when assessing the quality of a data item, dataset or other data product, a Data Quality Statement (DQS) is developed (see Appendix C for a DQS template). A DQS is a presentation of information about the quality of a data item or a collection of data items, using the Data Quality Framework. The purpose of quality statements is to clearly communicate key characteristics of the data which impact on quality, so that potential users can make informed decisions about fitness for use. Quality statements should report both the strengths and limitations of the data.

Quality statements vary in length and detail depending on the audience and medium for release. Staff members submitting data both internally and externally should prepare a DQS to accompany the submission. For regular submissions including automated electronic submissions a standing DQS can be used but should be reviewed regularly to ensure currency.

As a minimum a rating against ACT Health Data Quality Indicator should be reported (see ACT Health Data Quality Indicators below).

Application of the Data Quality Framework by producers of data

The focus on the fitness of data has emphasised the need to build quality into the production and delivery processes of data within the Directorate. It is recommended that producers of data consider the seven quality dimensions before designing collections, collecting data and producing outputs. This approach can enable informed decisions about factors including:

- appropriate methodology;

- desired outputs and their accessibility;
- the coherence of the collection in relation to other collections or products;
and
- the relevance of the collection given its purposes.

2.4.3 ACT Health Data Quality Indicators

Data Quality Indicators should be made available on all data quality documentation or with any contextual information for data or on reports themselves.

ACT Health considers that Interpretability and Accuracy are of equal importance. Interpretability can be measured by the extent metadata is available to clearly define, describe and explain the data. Accuracy can be measured by the extent and frequency that validation and error correction occurs against the data. It takes a combination of both clearly defined metadata and data quality processes to gain optimal clarity and value from the information.

ACT Health has developed the below matrix to assess data quality against Interpretability and Accuracy when providing reports or data. Statements against the other five dimensions should also be used to provide an overall assessment of the quality of data.

Data Quality Indicator

		Interpretability					
		Not registered	Registered	Some metadata defined	All metadata defined	All metadata standard	Credentialed
Accuracy	Unvalidated Data	0	1	2	3	4	5
	Some local edit checks	1	2	3	4	5	6
	All items edit checked	2	3	4	5	6	7
	Feedback reporting process	3	4	5	6	7	8
	Formal validation process including resolution	4	5	6	7	8	9
	Credentialed	5	6	7	8	9	10

Interpretability

Not Registered

- Report or Dataset is not registered in the eBIS Metadata Library

Registered

- Report or Dataset is registered in the eBIS Metadata Library
- None or very limited metadata has been defined

Some Metadata Defined

- Report or Dataset is registered
- Some data items have complete metadata defined OR All metadata items have some metadata defined

All Metadata Defined

- Report or Dataset is registered
- All metadata items have complete metadata defined

All Metadata Standard

- Report or Dataset is registered
- All metadata items have complete metadata defined
- All metadata items have been approved as standard

Credentialed

- Report or Dataset is registered
- All metadata items have complete metadata defined
- All metadata items have been approved as standard
- The report or dataset has been credentialed by the data governance and standards unit

Accuracy

Unvalidated Data

- No checks against the data have been made either by the source system or through error/edit reports

Some local edit checks

- Source system has some validation checks at data entry OR
- Error/edit reports are used to identify errors against some data items OR
- A combination of both source system checks and edit/error reports exist for some data items

All items edit checked

- All data items are checked either by the source system at data entry or through the use of error/edit reports
- Processes for resolution of errors exists

Feedback reporting process

- All data items are checked either by the source system at data entry or through the use of error/edit reports.
- Processes for resolution of errors exists
- Feedback is provided to individuals regarding data entry errors

Formal validation process including resolution

- All data items are checked either by the source system at data entry or through the use of error/edit reports.
- Documented processes for resolution of errors exists
- Documented feedback is provided to individuals regarding data entry errors
- Monitoring of data quality and resolution of issues is carried out

Credentialed

- All data items are checked either by the source system at data entry or through the use of error/edit reports.
- Documented processes for resolution of errors exists
- Documented feedback is provided to individuals regarding data entry errors
- Monitoring of data quality and resolution of issues is carried out
- The validation process has been credentialed by the data governance and standards unit

3 References

ⁱ Australian Council for Safety and Quality in Health Care 2001, in Victorian Quality Council 2003, *Better Quality, Better Healthcare: A safety and quality framework for Victorian health services*. Department of Human Services. Available from: www.health.vic.gov.au/qualitycouncil/pub/improve/framework.htm [cited 2013].

ⁱⁱ The Victorian Quality Council for Safety and Quality in Health June 2008, *A guide to using data for health care quality improvement*. Victorian Government Department of Human Services. Available from www.health.vic.gov.au/quality_council.

ⁱⁱⁱ Health Directorate *Information Management Framework 2012-2013*. ACT Government Health Directorate . Available from <http://inhealth/PandI/Shared%20Documents/Information%20Management%20Framework%20V1.pdf>.

^{iv} Taking the First Step with the PDCA (Plan-Do-Check-Act) Cycle. Available from www.bulsuk.com/2009/02/taking-first-step-with-pdca.html.

^v The DMAIC Model in Six Sigma. Available from www.sixsigmaonline.org/six-sigma-training-certification-information/articles/the-dmaic-model-in-six-sigma.html.

^{vi} 1520.0 - ABS Data Quality Framework, May 2009. Available from www.abs.gov.au/ausstats/abs@.nsf/Latestproducts/1520.0Main%20Features1May%202009?opendocument&tabname=Summary&prodno=1520.0&issue=May%202009&num=&view.

4 Appendices

{ Appendix A - Data Integrity Strategy }

{ Appendix B – Data Quality Policy }

Appendix C – ACT Health Data Quality Statement

Name of Submission / Report / Output:					
Date of Submission / Publication:					
Due Date of Submission:					
Name of Submitting Officer:					
Name of Business Unit:					
Scope of Data:					
Reference Period:					
Business Requirement / Alignment:					
Data Source Name(s):					
Data Source – Database Type:					
Data Format:					
Data Storage Location:					
Data extract procedures:					
Data processing procedures:					
Institutional Environment:					
Relevance:					
Timeliness:					
Accuracy:					
0. <input type="checkbox"/> Unvalidated Data	1. <input type="checkbox"/> Some local edit checks	2. <input type="checkbox"/> All items edit checked	3. <input type="checkbox"/> Feedback reporting process	4. <input type="checkbox"/> Formal validation process including resolution	5. <input type="checkbox"/> Credentialed
<i>Comment:</i>					
Coherence:					
Interpretability:					
0. <input type="checkbox"/> Not registered	1. <input type="checkbox"/> Registered	2. <input type="checkbox"/> Some metadata defined	3. <input type="checkbox"/> All metadata defined	4. <input type="checkbox"/> All metadata standard	5. <input type="checkbox"/> Credentialed
<i>Comment:</i>					
Accessibility:					

Data Quality Indicator:
(Accuracy + Interpretability)

--	--

Data Quality Indicator

		Interpretability					
		Not registered	Registered	Some metadata defined	All metadata defined	All metadata standard	Credentialed
Accuracy	Unvalidated Data	0	1	2	3	4	5
	Some local edit checks	1	2	3	4	5	6
	All items edit checked	2	3	4	5	6	7
	Feedback reporting process	3	4	5	6	7	8
	Formal validation process including resolution	4	5	6	7	8	9
	Credentialed	5	6	7	8	9	10

Interpretability

Not Registered	Report or Dataset is not registered in the eBIS Metadata Library.
Registered	Report or Dataset is registered in the eBIS Metadata Library; None or very limited metadata has been defined.
Some Metadata Defined	Report or Dataset is registered; Some data items have complete metadata defined OR All metadata items have some metadata defined.
All Metadata Defined	Report or Dataset is registered; All metadata items have complete metadata defined.
All Metadata Standard	Report or Dataset is registered; All metadata items have complete metadata defined; All metadata items have been approved as standard.
Credentialed	Report or Dataset is registered; All metadata items have complete metadata defined; All metadata items have been approved as standard; The report or dataset has been credentialed by the data governance and standards unit.

Accuracy

Unvalidated Data	No checks against the data have been made either by the source system or through error/edit reports.
Some local edit checks	Source system has some validation checks at data entry; OR Error/edit reports are used to identify errors against some data items; OR A combination of both source system checks and edit/error reports exist for some data items.
All items edit checked	All data items are checked either by the source system at data entry or through the use of error/edit reports; Processes for resolution of errors exist.
Feedback reporting process	All data items are checked either by the source system at data entry or through the use of error/edit reports; Processes for resolution of errors exists; Feedback is provided to individuals regarding data entry errors.
Formal validation process including resolution	All data items are checked either by the source system at data entry or through the use of error/edit reports; Documented processes for resolution of errors exists; Documented feedback is provided to individuals regarding data entry errors; Monitoring of data quality and resolution of issues is carried out.
Credentialed	All data items are checked either by the source system at data entry or through the use of error/edit reports; Documented processes for resolution of errors exists; Documented feedback is provided to individuals regarding data entry errors; Monitoring of data quality and resolution of issues is carried out; The validation process has been credentialed by the data governance and standards unit.