Data Quality Summary



**Instructions**

The Data Quality Summary[[1]](#footnote-1) is comprised of a series of questions that have been developed to enable the assessment of a data collection with respect to the five dimensions of data quality-relevance, accuracy, timeliness, coherence and interpretability.

The Data Quality Summary must be completed by the Data Custodian or nominated data quality staff at least annually (or at shorter intervals as deemed necessary by the Data Custodian). If data quality staff complete the Data Quality Summary, they must submit the completed document to the Data Custodian for their approval.

The Data Quality Summary is a minimum requirement and Data Custodians may wish to expand upon these reporting requirements. The findings of the Data Quality Summary must be made available to the Data Steward on request or as deemed appropriate by the Data Custodian.

**Data Quality Summary**

The Data Quality Summary is in four parts:

**Part A: Details about the data collection**

This part captures details about the data collection, including Data Steward and Data Custodian details.

**Part B: Assessment criteria**

This part captures information about the data collection with respect to the five dimensions of data quality. It comprises questions, as well as a section for comments, about each of the five dimensions of data quality to help determine whether data is fit for purpose. The five dimensions of data quality are not considered mutually exclusive, and the quality issues they raise may overlap. The importance of each dimension is not necessarily equal and may vary depending on the data, its context and user needs.

**Part C: Key findings**

Summary of key findings from the data quality assessment are to be documented in this part, including any strengths or limitations of the data collection.

**Part D: Document Approval**

This part is to be signed by the Data Custodian and details of any nominated data quality staff that have contributed to the Data Quality Summary documented.

**Other reporting requirements**

Where noteworthy limitations in data quality are identified, Data Custodians must ensure an Action Plan for quality improvement is documented no later than 3 months after completing the Data Quality Summary. Please refer to the Related Document - ‘Data Quality Improvement Action Plan’ for further information.

If you have any questions about this document please contact: Senior Policy Officer, Purchasing and System Performance Division. Email: [RoyalSt.PSPInfoManagement@health.wa.gov.au](mailto:RoyalSt.PSPInfoManagement@health.wa.gov.au).

# Part A - Details about the data collection

|  |  |
| --- | --- |
| ‘Full name of data collection’ - Enter here | |
| Data Steward | *Name, Position, Division, Location (i.e. Department/Name of Health Service Provider or Contracted Health Entity), Email.*  *Primary position responsible for data quality and influencing the organisational culture as it pertains to data quality.*  *Refer to the* [*Data Stewardship and Custodianship Policy*](file:///\\hdwa.health.wa.gov.au\shared\PAQ\EPG\IMR\DQ&G\Information%20Policy\2017\Data%20Quality\May%20-%20Revised\Data%20Stewardship%20and%20Custodianship%20Policy) *for further information.* |
| Data Custodian | *Name, Position, Division, (i.e. Department/Name of Health Service Provider or Contracted Health Entity), Email.*  *The position responsible for overseeing the reporting and monitoring requirements for data quality and improvement.*  *Refer to Data Stewardship and Custodianship Policy for further information.* |
| Name and type of data collection | *Includes name of data collection and the data type (e.g. administrative, registry, statutory collection, data warehouse).* |
| Brief description | *Brief outline of data collection and relevant information. Indicate whether the data collection is a local or enterprise data collection.*  *Refer to the* [*Data Stewardship and Custodianship Policy*](http://www.health.wa.gov.au/circularsnew/circular.cfm?Circ_ID=13318) *for further information.* |
| Purpose(s) | *Main stated purpose(s) of the data collection.* |
| Collection methodology | *Key features of the collection methodology (e.g. administrative or survey) and data collection method (e.g. administrative, data extracts, self-completion).* |
| Scope (coverage of relevant population) | *Population that is covered. May include geographic coverage, such as national, state or other, noting if there is adequate coverage for remote areas, non-English speakers and so forth.* |
| Frequency/timing | *Year(s) in which data have been collected, frequency of updates, releases.* |
| Size | *Number of records or sample size in most recent reference period.* |
| Further information | *If available (e.g. A web link with further information).* |

**Part B - Assessment criteria**

Listed below are statements relating to the five dimensions of data quality. At a minimum, please respond to each question by placing a ☑ (Yes) or a ☒ (No) in the check box provided.

If the statement is not applicable to your data collection leave the box blank ( Not applicable) and provide justification in the ‘Comments’ sections. You are also invited to leave comments about the strengths and limitations of the data in the ‘Comments' sections of the document.

# Relevance

The relevance dimension of data quality refers to the degree by which the data meets the needs of the user in terms of concepts measured and the population represented.

To assist in evaluating this dimension of data quality, please consider the following questions:

☐ Does the data correctly represent what it was designed to measure, monitor or report?

☐ Are the data items (or questions in surveys) collected of sufficient breadth and depth for their purpose?

☐ Is the data collection representative of the target population (e.g. survey data)?

☐ Are there any known gaps in the data? (e.g. non-responses, missing records, data not collected.)

Comments:

|  |
| --- |
|  |

# Accuracy

The accuracy dimension of data quality refers to the degree to which the data correctly describe the condition it was designed to measure.

To assist in evaluating this dimension of data quality, please consider the following questions:

☐ Does the data reflect the condition or situation it was designed to measure?

☐ Has the data been subject to quality assurance processes? (e.g. checking for errors at each stage of data collection and processing, or verifying data entry and making corrections if necessary.)

☐ Are potential or known sources of error or bias described?

☐ Is information available to help users evaluate the accuracy of the data and any level of error?

☐ Have there been any adjustments, changes or other factors that could impact the validity of data? (e.g. weighting, rounding, changes or flaws in data collection or verification methods).

Comments:

|  |
| --- |
|  |

# Timeliness

The timeliness dimension of data quality reflects the length of time between the availability of the data and the event or phenomenon it describes. Therefore, the timeliness dimension of data quality reflects the extent to which data are sufficiently up-to-date for a task.

To assist in evaluating this dimension of data quality, please consider the following questions:

☐ Are the data collected recorded or received as quickly as possible after the event or activity or for its intended use?

☐ Are the data collected, recorded or received the most current practically available for its intended use?

☐ Are the data up-to-date and current?

☐ Are the data available at regular intervals?

☐ Are there likely to be updates or revisions to the data after its release?

Comments:

|  |
| --- |
|  |

# Coherence

The coherence dimension of data quality refers to the internal consistency of a data collection, as well as how well it compares with other sources of information, within a broad analytic framework and over time.

To assist in evaluating this dimension of data quality, please consider the following questions:

☐ Have standard definitions, common concepts, classifications and data recording practices been used?

☐ Can data items (or questions in surveys) within the data collection be meaningfully compared?

☐ Are the data generally consistent with similar or related data collections?

☐ Are the data consistent with previous releases? (e.g. Have there been any significant changes in the way data items are defined, classified or counted over time? Have there been any changes in methodology or external impacts since the last data release?).

Comments:

|  |
| --- |
|  |

# Interpretability

The interpretability dimension of data quality refers to the ease with which the data can be understood by the user and the availability of information to help provide insight into the data.

To assist in evaluating this dimension of data quality, please consider the following questions:

☐ Is a data dictionary available to explain the meaning of data elements/items, format and relationships?

☐ Is information available about the data collection and methods of data collection?

☐ Is information available to help users correctly interpret the data and understand how it can be used?

☐ Is information available to explain ambiguous or technical terms used in the data?

Comments:

|  |
| --- |
|  |

**Part C – Key findings**

Based on the above information, summarise key findings, including any noteworthy strengths and limitations of the data collection with respect to the five dimensions of data quality.

Summary of findings:

|  |
| --- |
|  |

|  |
| --- |
| ☐ Noteworthy limitations in data quality were identified and will need to be addressed in a Data Quality Improvement Action Plan\*. |

\*Data Quality Improvement Action Plan must be completed no later than 3 months after the Data Custodian has approved the Data Quality Summary by signing Part D.

**Part D – Document Approval**

|  |  |  |  |
| --- | --- | --- | --- |
| Data Custodian details\* and approval. | | | |
| Name: | Click here to enter text. | **Position and Site Location:** | Click here to enter text. |
| Signature/ HE Number: |  | | |
| Date: | Select a date | | |

\*Should match Data Custodian details documented in Part A.

|  |  |  |  |
| --- | --- | --- | --- |
| Details of data quality staff member that completed or contributed to the Data Quality Summary (where appropriate). | | | |
| Name: | Click here to enter text. | **Position and Site Location:** | Click here to enter text. |
| Signature/  HE Number: |  | | |
| Date: | Select a date | | |

**This document can be made available in alternative formats   
on request for a person with a disability.**

© Department of Health 2017

Copyright to this material is vested in the State of Western Australia unless otherwise indicated. Apart from any fair dealing for the purposes of private study, research, criticism or review, as permitted under the provisions of the *Copyright Act 1968*, no part may be reproduced or re-used for any purposes whatsoever without written permission of the State of Western Australia.

1. The Data Quality Summary is based on the Australian Institute of Health and Welfare (2014) ‘[*Framework for Assessing Data Sources for Population Health Monitoring*](http://www.aihw.gov.au/WorkArea/DownloadAsset.aspx?id=60129548992)*’* and New South Wales Government (2015) [*Standard for Data Quality Reporting v1.2*](https://www.finance.nsw.gov.au/ict/sites/default/files/NSW%20Standard%20for%20Data%20Quality%20Reporting%20v1.2%20FINAL.pdf)though has been customised to make appropriate to the WA health system. [↑](#footnote-ref-1)