Health Information
and Standards

Background paper to support guidance for a data quality framework for health and social care

October 2018

Health Information and Quality Authority

An tÚdarás Um Fhaisnéis
agus Cáilíocht Sláinte

Safer Better Care

About the Health Information and Quality Authority

The Health Information and Quality Authority (HIQA) is an independent authority established to drive high quality and safe care for people using our health and social care services in Ireland. HIQA's role is to develop standards, inspect and review health and social care services and support informed decisions on how services are delivered.

HIQA aims to safeguard people and improve the safety and quality of health and social care services across its full range of functions.

HIQA's mandate to date extends across a specified range of public, private and voluntary sector services. Reporting to the Minister for Health and the Minister for Children and Youth Affairs, HIQA has statutory responsibility for:

- **Setting Standards for Health and Social Services —** Developing person-centred standards, based on evidence and best international practice, for health and social care services in Ireland.

- **Regulation —** Registering and inspecting designated centres.

- **Monitoring Children's Services —** Monitoring and inspecting children's social services.

- **Monitoring Healthcare Safety and Quality —** Monitoring the safety and quality of health services and investigating as necessary serious concerns about the health and welfare of people who use these services.

- **Health Technology Assessment —** Providing advice that enables the best outcome for people who use our health service and the best use of resources by evaluating the clinical effectiveness and cost-effectiveness of drugs, equipment, diagnostic techniques and health promotion and protection activities.

- **Health Information —** Advising on the efficient and secure collection and sharing of health information, setting standards, evaluating information resources and publishing information about the delivery and performance of Ireland's health and social care services.

Overview of the Health Information function of HIQA

Health is information-intensive, generating huge volumes of data every day. Health and social care workers spend a significant amount of their time handling information, collecting it, looking for it and storing it. It is therefore very important that information is managed in the most effective way possible in order to ensure a high-quality safe service.

Safe, reliable healthcare depends on access to, and the use of, information that is accurate, valid, timely, relevant and complete. For example, when giving a patient a drug, a nurse needs to be sure that they are administering the appropriate dose of the correct drug to the right patient and that the patient is not allergic to it. Similarly, lack of up-to-date information can lead to the unnecessary duplication of tests — if critical diagnostic results are missing or overlooked, tests have to be repeated unnecessarily and, at best, appropriate treatment is delayed or at worst not given.

In addition, health information has an important role to play in healthcare planning decisions — where to locate a new service, whether or not to introduce a new national screening programme and decisions on best value for money in health and social care provision.

Under section (8)(1)(k) of the Health Act 2007[1], the Health Information and Quality Authority (HIQA) has responsibility for setting standards for all aspects of health information and monitoring compliance with those standards. In addition, under section 8(1)(j), HIQA is charged with evaluating the quality of the information available on health and social care and making recommendations in relation to improving its quality and filling in gaps where information is needed but is not currently available.

Information and communications technology (ICT) has a critical role to play in ensuring that information to promote quality and safety in health and social care settings is available when and where it is required. For example, it can generate alerts in the event that a patient is prescribed medication to which they are allergic. Further to this, it can support a much faster, more reliable and safer referral system between the patient's GP and hospitals.

Although there are a number of examples of good practice, the current ICT infrastructure in Ireland's health and social care sector is highly fragmented, with major gaps and silos of information which prevent the safe, effective, transfer of information. This results in people using services being asked to provide the same information on multiple occasions.

In Ireland, information can also be lost, documentation is poor, and there is over-reliance on memory. Equally, those responsible for planning our services experience great difficulty in bringing together information in order to make informed decisions. Variability in practice leads to variability in outcomes and cost of care.

As a result of these deficiencies, there is a clear and pressing need to develop a coherent and integrated approach to improving the quality of health information, based on standards and international best practice. A robust health information environment will allow all stakeholders — patients and service users, health professionals, policy makers and the general public — to make choices or decisions based on the best available information.

Through its health information function, HIQA is working to support health and social care organisations in improving the quality of their data to better support the delivery, planning and monitoring of health and social care services.

Table of contents

1. Introduction

A considerable amount of data is collected on a regular basis about health and social care services in Ireland. Providers of health and social care services need quality data, not only at the point of service delivery but also at the point of decision-making in a format that is user friendly. Accurate, relevant and timely data is essential in order to improve health and social care, to inform decision-making, monitor diseases, organise services, support policy making, conduct quality research and plan for future health and social care needs, both at local and national level. Therefore, it is vital that there is trust in the quality of data and information produced in order to help provide safe and efficient health and social care to patients and service users.

Efforts to improve systems or processes within health and social care organisations must be driven by reliable data. Decisions are only as good as the information on which they are based. Where reliable data is available, organisations can accurately identify deficiencies, prioritise quality improvement initiatives and objectively assess whether change and improvement have occurred. To ensure that data is fit for purpose, health and social care organisations should adopt a systematic approach to assessing, improving and maintaining the quality of their data.

The importance of high-quality health and social care information in improving the safety and quality of patient care has been outlined in a number of key national strategy documents in recent years.[2,3,4] In particular, the *e-Health Strategy* published in 2013 highlights the importance of access to quality health data sources to drive improvements in the quality and safety of care.[5] Other recent developments that show the importance of quality health information include the establishment of the Office of the Chief Information Officer in 2015 and the publication of the Health Service Executive's *Knowledge & Information Strategy*.[6]

Based on international best practice, four key principles relating to health information have been identified which are based on maximising health gain for the individual and the population, specifically:

- health information is used to deliver and monitor safe and high-quality care for everyone
- health information should be of the highest quality and, where appropriate, collected as close as possible to the point of care
- health information should be collected once and used many times
- data collection should be 'fit for purpose' and cost-effective.

This background document outlines the international approaches to assessing and managing data quality and informed the development of HIQA's *Guidance on a Data Quality Framework for health and social care.*

1.1 What is quality?

Quality data refers to data that are 'fit for purpose' or 'fit for use'. A target for health and social care organisations is to produce data that are sufficiently accurate, timely and consistent to make appropriate and reliable decisions rather than aiming to produce perfect data.[7]

Quality health and social care is dependent on the access to and use of quality data. Data users can have greater confidence in the data that they obtain where a systematic approach is taken to assessing the quality of that data.

The benefits of quality data include:

- facilitation of better informed strategic, national and local planning for health and social care
- provision of safer, better quality care for patients
- better informed national and local policies
- greater contributions to research
- improved population health.

Internationally, there are seven recognised essentials for improving data quality in health and social care and, subsequently, the overall quality of care. These are:

- leadership and management
- policies and procedures
- standardisation
- training
- make data available
- data quality audits
- dimensions of data quality.[8]

The quality of data can be determined through assessment against a number of key dimensions. Data quality dimensions are 'a set of data quality attributes that represent a single aspect or construct of data quality'.[9] This background document highlights the numerous dimensions of data quality, each with different interpretations, which have been identified in the literature to describe data quality. Table 2 details the most commonly used dimensions internationally, namely, relevance, accuracy, reliability, timeliness, punctuality, coherence, comparability, accessibility and clarity.[10,11,12] HIQA has adopted these international dimensions of quality, which are recognised as international best practice in the assessment of data quality within the health and social care sector.

By assessing data quality it is possible to:

- establish a baseline for data quality and identify any areas for improvement
- assess the impact of any changes in practice, policies or procedures on data quality
- report on improvements in data quality.

To ensure that data is fit for purpose, organisations should adopt a systematic approach to assessing the quality of their data.

1.2 Purpose and scope of this background paper

The objective of this background document is to review national and international best practice in relation to data quality frameworks and the dimensions of data quality that are utilized in order to systematically assess, monitor, evaluate and improve the quality of their data and information. This background document has been prepared to inform HIQA's work on the development of *Guidance on a Data Quality Framework for health and social care* in Ireland. As health and social care organisations hold a rich source of data, this background paper aims to outline for organisations, best practice in relation to data quality assessment internationally, identifying processes that should be undertaken to ensure that data is 'fit for purpose'.

A **data quality framework** outlines the approach and provides the tools that can be used by organisations to systematically assess, document and improve data quality. The components of a data quality framework include the following:

- data quality strategy
- data quality assessment methodology
- reporting on data quality
- data quality improvement cycle.

1.3 Methodology

The focus of this background paper is to determine current practices internationally in relation to the use of data quality frameworks, the components they include and the dimensions under which data quality is assessed by international organisations. In line with the guidance development process, HIQA undertook a detailed desktop review to identify examples of best practice internationally. The review identified the frameworks outlined in Table 1, as suitable for inclusion in this review. The frameworks are a mixture of data quality frameworks and quality assurance strategies/assessment frameworks from statistical bodies.

Table 1. International frameworks reviewed to support *Guidance on a Data Quality Framework for health and social care*

CIHI's Information Quality Framework [13]
Statistics Canada Quality Assurance Framework, Third Edition 2017[14]
ACT Health Data Quality Framework[15]
ABS Data Quality Framework[16]
Quality Assurance Framework of the European Statistical System[11]
United Nations Statistics Division, Guidelines for a Template for a National Quality Assurance Framework (NQAF)[12]
Quality framework and Guidelines for OECD statistical activities[17]
GS1 Australia Healthcare Data Quality Framework Version 1.0[18]
South African Statistical Quality Assessment Framework (SASQAF)[19]
NHS Digital Data Quality Assurance Strategy 2015-2020[20]
Independent Hospital Pricing Authority (IHPA) Data Quality Assurance Framework[21]

The findings of this review are presented within this background document, with a chapter dedicated to each major framework reviewed.

2. Data quality in national health and social care data collections in Ireland

2.1 HIQA's work in this area to date

HIQA previously published a number of key documents relating to data quality, these include:

- *International Review of Data Quality*, (2011)[22]
- *What you should know about Data Quality: a guide for health and social care staff*, (2012)[8]
- *International review of health and social care data collections* (2016).[23]

In 2011, HIQA undertook a survey of national health and social care data collections in Ireland. This work was undertaken to establish themes to support the development of the *Information Management Standards for National Health and Social Care Data Collections*.[24] The survey found that while data quality was a focus within all national data collections, quality improvement programmes were limited, as was high level organisational responsibility for data quality. The fragmented nature of national data collections in Ireland meant that linkages of data sources were, and still are, difficult and costly to achieve. Greater interoperability would allow national data collections to share data more readily and reduce duplication of effort in recording similar information. The survey also showed a lack of a strategic approach to reporting on data quality across most national data collections.[25]

In 2017, HIQA published *Information management standards for national health and social care data collections.*[24] The purpose of these standards is to provide a roadmap to improve the quality of national health data and information, which will ultimately contribute to the delivery of safer, more reliable health and social care for service users and patients. Data is specifically addressed in Standard 6 under the theme *Use of Information.* In this standard, it is recommended that national data collections develop a data quality framework to assist them in systematically assessing, monitoring, evaluating and improving the quality of their data.

HIQA developed a structured review programme for assessing compliance with the Information Management Standards, which began in 2017. *The Guide to the Health Information and Quality Authority's review of information management practices in national health and social care data collections* outlines the aims and objectives of the review programme in detail.[26] The aim of this new programme of work is to:

- determine the information management practices of the individual national data collections
- identify areas of good practice and areas where improvements are necessary
- work with the national data collections to achieve compliance with the standards, ultimately improving the quality and maximising the use of national health information.

2.2 Examples of data quality reports within Irish health and social care organisations

Quality reporting is the preparation and dissemination, on a regular or irregular basis, of reports conveying information about the quality of a statistical product or survey. A quality report provides information on the main quality characteristics of a product so that the user should be able to assess product quality. In the optimal case, quality reports are based on quality indicators.[7] Examples of data quality reports by national health and social care data collections include:

- Standard reports on methods and quality from the Central Statistics Office (CSO)[27]
- *Data Quality and Completeness at the Irish National Cancer Registry* from the National Cancer Registry of Ireland (NCRI).[28]

2.2.1 CSO's standard reports on methods and quality[27]

The CSO's standard reports on methods and quality are a series of reports that provide information to users on the methodology in place in the particular survey area examined together with an assessment of the quality of the resulting survey output. The reports contain general information about the survey area, statistical concepts and methods applied, details of the production of the statistics, data processing and quality assurance techniques applied and an overview of the quality of the statistics under the following dimensions of data quality[27]:

- relevance
- accuracy and reliability
- timeliness and punctuality
- accessibility and clarity
- comparability
- coherence.

Standard reports on methods and quality are in the process of being completed for all CSO releases and publications and will be made available on the website as they are finalised. The following are a sample of some of the reports available:[27]

- Census of Population (last edited 2009)
- the Quarterly National Household Survey (last edited 2017)
- Deaths registration (last edited 2015)
- Births registration (last edited 2015).

2.2.2 Data Quality and Completeness at the Irish National Cancer Registry[28]

This study, published in 2012, aims to give a broad overview of data quality at the NCRI by examining the comparability, completeness and validity of the data using some of the techniques outlined in papers by Bray and Parkin.[29,30]

To evaluate comparability, it is important that information collected by the registry is coded and classified according to international guidelines so that Irish cancer statistics can be compared to other countries worldwide. Issues considered in this area include International Classification of Diseases (ICD) codes, definitions of incidence and how multiple tumours are handled.

Completeness is a measure of the extent to which the NCRI is capturing diagnosed cancers in Ireland. There are various ways to examine this, and the methods fall broadly into two main categories: semi-quantitative or quantitative.

Validity in this case refers to whether a case has been historically verified. If the only information regarding a case is a death certificate, then, due to lack of follow-up information, it cannot be confirmed that it was a cancer case. Therefore, a low level of "death certificate only" is considered an important maker for checking the validity of the data. Another important indicator of validity used in this case was the percentage of cases morphologically verified or listed as "primary site unknown".

The overall findings of the report suggest that the comparability, completeness and validity of the data at the NCRI are of a good standard. The report found that the use of international codes within the NCRI helped to facilitate comparison to rates of cancers in other countries. The completeness of the registry was found to be high, with the completeness found to range between 97–99%, depending on the methods used. The rate of "death certificate only" cases was found to be low at 1.3%, highlighting the validity of the data (this is at the median level for the European registries).

3. Summary of international evidence

In order to assess, document and improve data quality, a framework is required to provide appropriate, standardised methods for improving data quality. All of the frameworks included in this report, as detailed in Table 1, provide a definition for data quality and a means of assessing data quality to ensure it is fit for purpose. To compare and contrast the different frameworks, this report focuses on the following aspects of each framework:

- the dimensions of data quality used
- the components of the framework
- how the framework is used in practice.

3.1 Data quality dimensions

There is no single accepted definition of data quality; this may be due to the varying focuses of the data quality frameworks. The frameworks have various functions and the types of data being considered can vary widely.[31]

The Canadian Institute for Health Information (CIHI) recently updated its dimensions of data quality to align to the dimensions used by the United Nations. These dimensions are also used by the European Statistical System (ESS), Eurostat and Statistics Canada, and they closely align to those used in the CSO Standard Reports on Methods and Quality.[13] These dimensions and their definitions are detailed in Table 2.

Table 2. Dimensions of data quality

Dimension	Definition
Relevance	Relevant data meets the current and potential future needs of users.
Accuracy and reliability	The accuracy of data refers to how closely the data correctly describes what it was designed to measure. Reliability refers to whether that data consistently measures, over time, the reality that it was designed to represent.
Timeliness and punctuality	Timely data is collected within a reasonable agreed time period after the activity that it measures. Punctuality refers to whether data are delivered or reported on the dates promised, advertised or announced.
Coherence and comparability	Coherent and comparable data is consistent over time and across providers and can be easily combined with other sources.
Accessibility and clarity	Data are easily obtainable and clearly presented in a way that can be understood.

All of the frameworks reviewed apply defined dimensions or metrics to assess data quality. While each framework defines a set of dimensions of data quality, international agreement has not been reached on one defined set of dimensions. Furthermore, where organisations use the same dimensions of data quality, there are often discrepancies in the definitions of that dimension due to the contextual nature of the definitions.[32] While some dimensions are consistent across most of the frameworks, others are not consistently used. Table 3 highlights the data quality dimensions used by each organisation.

In light of the most recent international evidence detailed in this report, the dimensions of data quality used by HIQA have been updated to match those detailed in Table 2, bringing HIQA's data quality practices in line with international best practice.

Tables detailing the varying definitions used of the dimensions of data quality can be found in Appendix 2

3.2 Key components of a data quality framework

Common components of the frameworks reviewed include:

1. A data quality strategy — an organisation's formalised approach to address data quality which sets out the activities that an organisation needs to undertake in order to strengthen their approach to the collection, handling, use and dissemination of data and information.

2. A data quality assessment tool — a set of criteria to comprehensively assess data sources across the five dimensions of quality, as detailed in Part 2 of this guidance.

3. Reporting on data quality — data quality reports can include internal or external data quality assessment reports, reporting on key performance indicators or metrics and producing "data quality statements".

4. A data quality improvement cycle — a description of how continuous improvement of data quality is assured by the health or social care organisation. It encompasses the processes and methodologies applied by organisations as part of their data quality improvement initiatives.

3.2.1 Data quality strategy

A data quality strategy should provide a roadmap of how a health or social care organisation manages and continuously improves the quality of their data. It formalises the approaches for identifying, documenting and ensuring the implementation of data quality practices.[33] A data quality strategy can be a standalone document or can be incorporated within an overall information management strategy. Information on the following may be included as part of a data quality strategy:

- governance, leadership and management structures
- policies and procedures

- training, education and development programme
- data quality audit
- standards.

3.2.2 Data quality assessment tools

Data quality assessment tools such as those used by CIHI and Australian Capital Territory (ACT) Health enable each organisation, data users and data providers to assess the quality of the data and thus determine its fitness for use. CIHI's 2017 data source assessment tool provides a set of criteria that allow for data source quality assessment across the five dimensions of data quality, allowing for appropriate determination of a given set of data's fitness for use.[34] The ACT Health data quality assessment tool divides the dimensions of data quality into characteristics that are used systematically to assess the fitness for purpose of the data.[15]

3.2.3 Reporting on data quality

Reporting on data quality is another important component of a data quality framework. Reporting on data quality can include internal or external data quality assessment reports, reporting on key performance indicators and data quality statements. Data quality reports communicate to the data user how "fit for purpose" the data is.

3.2.4 Data quality cycles

Data quality cycles outline a continuous approach the health and social care organisations can undertake to improve the quality of their data. Two such data quality cycles are detailed in *CIHI's Information Quality Framework* and *ACT Health Data Quality Framework*.[13,15]

CIHI's quality activities include the following stages:

- Prevent — preventing quality issues before they occur, which includes the use of standards, vendor specifications, conducting training and client support and system audits and edits.
- Monitor and control — activities undertaken as data is submitted and analysed. This includes error reports and corrections of data as it comes in to CIHI, issue management through submission of reports to data stewards and advanced surveillance methodologies which encompass sophisticated statistical techniques to flag anomalies.
- Evaluate and document — once a data set of information product is considered final, various activities come into effect to evaluate and document overall quality of the data. This includes documentation and metadata; quality assessments; stakeholder consultations; indicators and reports; and validation, reabstraction studies.
- Improve — improvement to data and information occurs iteratively over the course of conducting quality activities. The continuous improvement cycle results in high-quality data and information and improves prevention, monitoring and evaluation efforts. Examples of CIHI's improvement activities include collaboration with

stakeholders; system changes; and updates to standards, processes, resources and trainings.

ACT Health use a data quality improvement cycle which goes through the following steps[15]:

- **Recognise** — actively identifying data issues that need to be addressed though 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.

3.3 How data quality frameworks are used in practice

Data quality frameworks provide an overarching structure for an organisation's data quality practices. Many of the frameworks included in this background document have been developed by organisations for their own internal use, for example, CIHI's information quality framework has been developed for use in all its data holdings,[13] Statistics Canada's quality assurance framework has been developed for quality management within the organisation,[14,15] and the ACT Health data quality framework was developed to improve the quality of data created and managed by ACT Heath.[14,15] Other frameworks have been developed in a more open and flexible manner to be used by a wider group of stakeholders: the Australian Bureau of Statistics (ABS) data quality framework is designed for a wide range of users including government and statistical agencies[16] and the ESS quality assurance framework and the UNSTATS National Quality Assurance Framework were both developed for general use by statistical agencies at a European and national level.[11,12] Table 4 provides an summary of frameworks used internationally.

Table 3. Data quality dimensions used in international frameworks[*]

Dimension	CIHI	Statistics Canada	ACT Health	ABS	ESS	UNSTATS	OECD	GS1 Australia	SASQAF
Accessibility	✓	✓	✓	✓	✓	✓	✓		✓
Accuracy	✓	✓	✓	✓	✓	✓	✓	✓	✓
Clarity	✓	✓			✓	✓			
Coherence	✓	✓	✓	✓	✓	✓	✓		✓
Comparability	✓	✓			✓	✓			✓
Completeness								✓	
Consistency								✓	
Credibility							✓		
Institutional environment			✓	✓					
Integrity									✓
Interpretability		✓	✓	✓			✓		✓
Legibility									
Managing metadata		✓				✓			
Methodological soundness									✓
Prerequisites of data quality					✓	✓			✓
Punctuality	✓	✓			✓	✓			
Relevance	✓	✓	✓	✓	✓	✓	✓		✓
Reliability	✓	✓			✓	✓			
Standards based								✓	
Timeliness	✓	✓	✓	✓	✓	✓	✓	✓	✓
Usability									
Validity									

[*] The NHS Digital Data Quality Assurance Strategy and the IHPA Data Quality Assurance Framework have not been included as they do not detail dimensions of data quality

Table 4. **Summary of international evidence**[†]

Framework	Framework components	Dimensions of data quality	Use of the framework
CIHI's Information Quality Framework (13)	The Information Quality Framework consists of the following four key elements: ▪ Foundation — strategies and principles, policies and procedures, corporate tools, standards, best practice guidelines, knowledge exchange, culture ▪ Activities — prevent, monitor and control, evaluate and document, improve ▪ Outputs — data/information quality reports and indicators, improvement action plans, metadata ▪ Outcomes — increased knowledge, increased trust, increased use, increased quality.	▪ relevance ▪ timeliness and punctuality ▪ comparability and coherence ▪ accessibility and clarity ▪ accuracy and reliability	CIHI's Information Quality Framework provides an overarching structure for CIHI's quality management practices related to capturing and processing data and transforming it into information products. The Data Source Assessment Tool, which facilitates the 'activities' element of the Information Quality Framework, can be applied to many different types of data sources and is comprehensive enough to meet the quality information needs of data users while supporting remit of data stewards in CIHI. The criteria within the tool can be adapted to meet the specific needs of an organisation.
Statistics Canada Quality	The QAF consists of twelve stand-alone chapters, each defined by a quality management theme. All	The dimensions listed under "statistical outputs" most closely align to the dimensions	The quality assurance framework serves as the highest-level governance tool for quality management at Statistics Canada. It describes

---

[†] The NHS Digital Data Quality Assurance Strategy and the IHPA Data Quality Assurance Framework have not been included as they do not detail dimensions of data quality

Assurance Framework (QAF), Third Edition 2017[14]	chapters share a common three-part approach, divided into description, assessment and implementation.	of data quality used in data quality frameworks. These focus on: ▪ relevance ▪ accuracy and reliability ▪ timeliness and punctuality ▪ accessibility and clarity ▪ coherence and comparability ▪ interpretability and management of metadata.	the measures that the agency has to put in place to manage quality and provides guidance to statistical programme areas as they develop and implement quality management strategies to meet their users' needs. It may serve as a useful reference for those interested in the production of official statistics.
ACT Health Data Quality Framework[15]	The framework comprises of the following components: ▪ the Data Integrity Strategy ▪ the Data Quality Policy ▪ the Data Quality Improvement Cycle ▪ the Data Quality Assessment Tool.	▪ accessibility ▪ accuracy ▪ coherence ▪ institutional environment ▪ interpretability ▪ relevance ▪ timeliness.	The ACT Health Data Quality Framework provides an objective approach to assessing and improving the quality of data created and managed by ACT Health. It 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 the analysis of the fitness for purpose of a data collection.
Australian Bureau of Statistics (ABS) Data Quality	The ABS DQF is comprised of seven dimensions of quality, reflecting a broad and inclusive approach to quality definition and assessment.	▪ accessibility ▪ accuracy ▪ coherence ▪ institutional environment	The ABS DQF is designed for use by a range of data users and providers in different settings, including government agencies, statistical agencies and independent research agencies. It is a general framework to enable a comprehensive

Framework (DQF)	The framework provides: - an explanation of each of the seven dimensions of data quality - questions to be considered for the purpose of assessing the dimension.	- interpretability - relevance - timeliness.	and multi-dimensional assessment of the quality of a statistical dataset, product or release.
Quality Assurance Framework of the European Statistical System[11]	The ESS Quality Assurance Framework identifies activities, methods and tools that facilitate the practical and effective implementation of the European Statistics Code of Practice indicators. The Quality Assurance Framework is divided into three key themes: **Institutional environment:** - commitment to quality. **Statistical processes:** - sound methodology - appropriate statistical procedures - non-excessive burden on respondents - cost effectiveness.	The dimensions listed under "statistical outputs" most closely align to the dimensions of data quality used in data quality frameworks. These focus on: - relevance - accuracy and reliability - timeliness and punctuality - coherence and comparability - accessibility and clarity	The ESS Quality Assurance Framework is a guiding tool to assist the implementation of the European Statistics Code of Practice by statistical authorities at national (including other national data producers) and European levels, thereby becoming an important instrument of the ESS. The open and flexible nature of the ESS Quality Assurance Framework allows the specific selection of recommended activities, methods and tools that better fit the context of a specific statistical authority.

	Statistical output: ▪ relevance ▪ accuracy and reliability ▪ timeliness and punctuality ▪ coherence and comparability ▪ accessibility and clarity. A set of indicators accompany each of the principles, demonstrating how compliance can be achieved.		
United Nations Statistics Division, Guidelines for a Template for a National Quality Assurance Framework (NQAF) [12]	The Template for a Generic National Quality Assurance Framework includes the following components: ▪ quality context ▪ quality concepts and frameworks ▪ quality assurance guidelines ▪ quality assessment and reporting ▪ quality and other management frameworks.	The dimensions listed under "managing statistical outputs" most closely align to the dimensions of data quality used in data quality frameworks. These focus on: ▪ assuring accessibility and clarity ▪ assuring accuracy and reliability ▪ assuring coherence and comparability ▪ assuring relevance ▪ assuring timeliness and punctuality	The template is intended to be used as a tool to provide the general structure within which countries that choose to do so can formulate and operationalise national quality frameworks of their own or further enhance existing ones.

		▪ managing metadata.	
Quality framework and Guidelines for OECD statistical activities[17]	The Organisation for Economic Co-operation and Development (OECD) Quality Framework has four elements: ▪ a definition of quality and its dimensions ▪ a procedure for assuring the quality of proposed new statistical activities ▪ a procedure for evaluating the quality of existing statistical activities on a regular basis ▪ a set of broad principles on which OECD statistical activities are to be conducted and quality guidelines covering all phases of the statistical production process.	▪ accessibility ▪ accuracy ▪ coherence. ▪ credibility ▪ interpretability ▪ relevance ▪ timeliness.	The OECD Quality Framework outlined in this document focuses on improving the quality of data collected, compiled and disseminated by the OECD through an improvement in the Organisation's internal statistical processes and management, though there will also be a positive knock-on effect on the quality of data compiled at the national level.
GS1 Australia Healthcare Data Quality Framework Version 1.0[18]	Within the GS1 Australia Healthcare Data Quality Framework, there are five phases: ▪ Phase A: Action National Product Catalogue (NPC) Data Quality Report — Commit to the Data Quality Framework. Action NPC Data Quality Report- begin data	GS1 defines data quality to encompass the following pillars: ▪ completeness ▪ standards-based ▪ consistency ▪ accuracy ▪ time-stamped.	The framework has been established to assist all stakeholders in the Australian healthcare sector who supply or are a touch-point of product master data.

	remediation. ▪ Phase B: Conduct Data Quality Self-Assessment — Manage and Complete Data Quality Self-Assessment using scorecard. Report results (internally). ▪ Phase C: Plan Data Quality Management system — Plan Data Quality Management (DQM) based on self-assessment. ▪ Phase D: Document standard operating procedures (SOPs) and Implement Data Quality Management system processes. Report on data quality outcomes (internally). ▪ Phase E: Review and Embed Data Quality Management — Review and adjust SOPs, embed DQMS into Policy and finalise SOPs and ensure DQMS authorities and succession is in place. The final element of the framework is to ensure sustained data quality. This		

	involves conducting a second data quality (DQ) self-assessment to ensure that internal improvements have occurred. The NPC or internal DQ reports need to be monitored and the necessary remediation needs to be actioned.		
South African Statistical Quality Assessment Framework (SASQAF)[19]	The framework covers the as quality aspects of the entire statistical value chain and certifies national statistics on one of four levels: ▪ Level four: quality statistics ▪ Level three: acceptable statistics ▪ Level two: questionable statistics ▪ Level one: poor statistics. The quality dimensions set out descriptions and key components, the related indicators for each dimension, standards that need to be adhered to in relation to a quality indicator, and provide benchmarks for the standards which determine the four quality levels as detailed.	▪ accessibility ▪ accuracy ▪ comparability and coherence ▪ integrity ▪ interpretability ▪ methodological soundness ▪ relevance ▪ timeliness.	The South African Statistical Quality Assessment Framework (SASQAF) provides a framework and criteria used for evaluating and certifying statistics produced by government departments and other organs of state and, in some circumstances, by non-governmental institutions and organisations. The main purpose of the South African Statistical Quality Assessment Framework is to provide a flexible structure for the assessment of statistical products. The South African Statistical Quality Assessment Framework can be used for: ▪ self-assessment by producers of statistics ▪ reviews performed by a data quality assessment team ▪ assessment by data users, based on the producing agency's quality declaration ▪ assessment by international agencies, based on the quality declaration.

4. Canadian Institute for Health Information (CIHI), Information quality framework, 2017

4.1 CIHI

CIHI is an independent, not-for-profit organisation that provides essential information on Canada's health systems and the health of Canadians. It was established in 1994 to forge a common approach to Canadian health information. It aims to provide comparable and actionable data and information that are used to accelerate improvements in healthcare, health system performance and population health across Canada. To date, CIHI is responsible for 28 pan-Canadian databases across health sectors. Stakeholders use its broad range of health system databases, measurements and standards, together with evidence-based reports and analyses, in their decision-making processes.[35]

4.1.1 Data quality in CIHI

Quality is at the heart of everything that the CIHI does. CIHI's data and information quality programme is recognised nationally and internationally for its comprehensiveness and high standards.[36] CIHI's strategic plan 2016–2021[37] is focused on providing stakeholders with the information that they require. A specific goal of the plan is for CIHI to be a trusted source of standards and quality data. Thus, CIHI strives to continuously improve data and information quality within CIHI and the broader health sector, with data quality being a key corporate priority, fundamental to CIHI's mandate. While direct responsibilities for data quality fall to the jurisdictions, CIHI recognises that improving data and information quality is a collaborative effort and works with its data suppliers to support improvement activities. This collaborative approach is designed to meet the changing and expanding user requirements and expectations of CIHI's data collections.[38]

CIHI regularly performs a series of activities to help monitor data quality and prevent and detect data quality issues. These include:

- the programme area responsible for the data holding completes a full assessment using the data quality framework assessment tool at least once every three years or if the holding undergoes significant change
- providing educational sessions to data providers on data requirements and data quality issues
- offering a coding query service to health information management professionals and other data collectors
- building systems that automatically check for data quality issues and providing feedback to data suppliers
- assessing and documenting the quality of data and providing this information to suppliers and users of the data

- preparing annual jurisdictional reports on the status of data quality for provincial and territorial deputy ministers
- reabstraction studies which are used to evaluate clinical data in acute care.

CIHI uses the term "fit for use" to define data quality, in line with international standards and best practices. [39]

4.1.2 CIHI's Information Quality Plan

In 2017, CIHI published an Information Quality Plan.[39] Building on their existing strong quality foundation, this Information Quality Plan supports the achievement of CIHI's overall strategic goals. The plan aims to:

- ensure that CIHI continues to be a trusted source of health information
- ensure that CIHI's information is for and meeting the broadening needs of stakeholders
- ensure that quality improvement initiatives are focused in the right areas
- evolve and improve the effectiveness and efficiency of CIHI's quality management practices.

The focus of the plan is information quality and looks at the entire life cycle of creating and acquiring data, processing it and transforming it into information products. Information quality encompasses the data quality of the data sources and also incorporates how the information is presented and accessed and can be understood. It encompasses how information fits together to paint a coherent picture and the quality management of all the processes involved across the life cycle.

The information quality plan consists of three main objectives:

Objective 1: Develop and launch an information quality framework that brings together and enhances CIHI's quality management practices.

Common approaches to information quality management need to be developed and best practices more easily shared. CIHI's commitment to quality needs to be continually reinforced. Quality is a shared responsibility across the organisation, and every employee needs to understand his or her role. CIHI also needs to support quality stewardship among many other organisations and people involved in supplying data to CIHI. Examples of initiatives related to this objective include:

- evolving and standardising data evolution tools and processes by developing efficient and effective procedures and tools for evaluating different types of data
- reinforcing CIHI's quality culture by enhancing internal and external knowledge sharing opportunities and by developing standard internal and external education.

Objective 2: Improve the accuracy, comparability and coherence of CIHI's data and information.

There is a need to combine information from multiple data holdings to produce more valuable-system level indicators and information. Identifying issues and providing relevant,

timely information to data providers about the quality of their data are key processes in maintaining and improving overall quality. Examples of initiatives related to this objective include:

- streamlining the internal processes to identify and manage quality issues by implementing a corporate tool to tract and manage operational issues and by creating a corporate action plan for high-quality reporting to data providers
- improving and streamlining data quality reporting to data providers by enhancing provincial and territorial data quality reports and by streamlining and automating routine data quality monitoring and dissemination of feedback to data providers.

Objective 3: Improve the information (metadata) about CIHI's data sources and information products.

Data providers need easy access to reference materials and metadata such as user manuals, standards and data dictionaries to ensure high-quality collection and submission of data. Data users need access to reference materials and metadata to ensure that they use the data and information appropriately. Transparency around the quality of data sources and methodologies used to create the information will build trust and prompt information quality improvement. Examples of initiatives related to this objective include:

- improving the consistency and availability of metadata on CIHI's data holdings, indicators, statistics and information products
- publishing provincial and territorial data quality indicators and metrics on CIHI's website and developing core data quality indicators for all other data holdings
- improving the transparency and consistency of quality processes, standards and methodologies by developing a consistent approach to both the disclosure of data limitations and the documentation of revisions to data and information.

4.2 CIHI's Information Quality Framework, 2017

CIHI's Information Quality Framework[13] provides an overarching structure for CIHI's quality management practices relating to capturing and processing data and transforming it into information products. In 2017, CIHI explicitly moved from the 2009 Data Quality Framework to the current Information Quality Framework. This change was brought about in order to ensure the quality of both data coming in and also to ensure quality in how it is processed and transformed into information products. The information quality framework looks at the entire life cycle of creating and acquiring data, processing it and transforming it into information products. Information quality encompasses the data quality of the data sources and also incorporates how the information is presented and accessed; how it can be understood; how information fits together to paint a coherent picture; and the quality management of all the processes involved across the cycle.[39]

The Information Quality Framework is a high level framework, describing how data is transformed into information across five phases of the information life cycle: capture,

submit, process, analyse and disseminate. Figure 1 signifies CIHI's quality management practices. The framework has four components:

- the foundation, which includes some of the elements necessary to achieve quality, for example, strategy and principles, policies and procedures, corporate tools, standards, best practice guidelines, knowledge exchange and culture
- the activities, including prevention, monitoring and control, evaluation and documentation, and improvement
- the outputs, including data/information quality reports and indicators, improvement action plans and metadata
- the outcomes, including improved knowledge, increased trust, increased use of information which is of better quality.

One key change since the 2009 Data Quality Framework is that CIHI have updated their dimensions of data quality to align with those used by the United Nations National Quality Assurance Framework.[40] These dimensions are as follows:

- relevance
- accuracy and reliability
- timeliness and punctuality
- coherence and comparability
- accessibility and clarity.

Figure 1. CIHI's Information Quality Framework

Information Quality Framework

Data — Information

| Capture | Submit | Process | Analyze | Disseminate |

Foundation
- Strategy and principles
- Policies and procedures
- Corporate tools
- Standards
- Best practice guidelines
- Knowledge exchange
- Culture

Activities
- Prevent
- Monitor and control
- Evaluate and document
- Improve

Outputs
- Data/information quality reports and indicators
- Improvement action plans
- Metadata

Outcomes
- Increased knowledge
- Increased trust
- Increased use
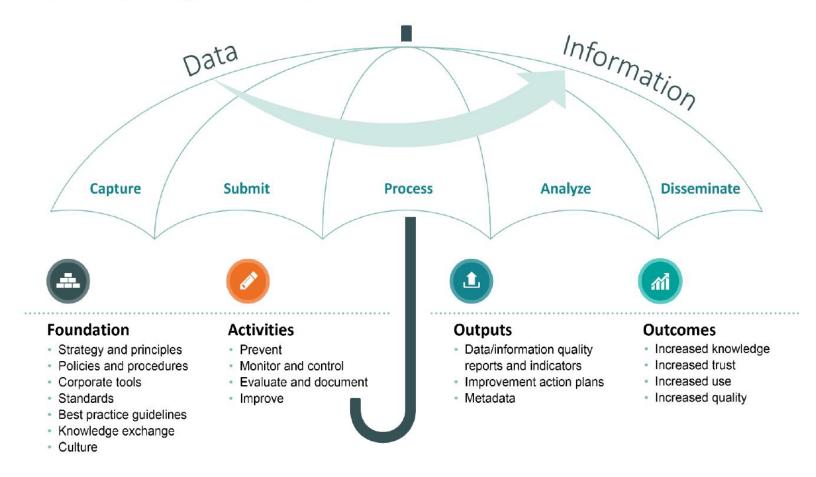- Increased quality

Figure 2. CIHI's Quality Activities

4.3.1 CIHI's quality principles

CIHI's quality activities, as detailed in Figure 2, are guided by a set of principles that acknowledge CIHI's role in achieving information quality, in conjunction with the other key stakeholders that are part of the information supply chain.[13] The quality principles are:

- quality is a multi-dimensional concept
- quality is relative and based on users' needs
- users must be informed about the quality of data or information so that they can assess its fitness for their use
- quality is a responsibility that is shared by everyone involved in the information supply chain
- quality assurance occurs at every stage of the information life cycle, with a focus on issue prevention
- quality is a feature of organisational culture
- standards are a foundation for quality information.

4.3.2 CIHI's information life cycle

CIHI's information life cycle is adapted from the high-level processes of the General Statistical Business Process Model (GSBPM) and describes the following steps for turning data into information.

- **Capture —** CIHI's influence over quality at this stage consists of producing standards and supporting providers in completing and enhancing their data capture.
- **Submit —** CIHI supports high-quality data submission through the use of submission specifications, standards, tools, schedules, education and client support.
- **Process —** both CIHI and data providers are involved in iteratively cleaning and correcting any issues identified during the processing phase.
- **Analyze —** CIHI works with external stakeholders to ensure the information products will meet their needs; stakeholders also provide vital expertise and advise on the content and methods used.
- **Disseminate —** this stage includes activities leading to the release, communication and promotion of results. It includes activities related to supporting users' understanding of the information and collaborating with stakeholders to increase their capacity to use data and analysis.[13]

CIHI have also developed a new Data Quality Assessment Tool.[34] This replaces the Data Quality Framework Assessment Tool in the 2009 Data Quality Framework. The current version is specifically for use internally in CIHI; it provides a set of criteria that allow for data source quality assessment across CIHI's five dimensions of quality, allowing for appropriate determination of a given set of data's fitness for use. An external version is in development which will provide criteria definitions.

4.3 CIHI's Data Source Assessment Tool

The Data Source Assessment Tool provides a set of criteria that enable CIHI to comprehensively assess its data sources across five dimensions of data quality and to determine the fitness for use of a particular dataset.[34] The tool can be applied to many different data sources and can be used to create quality benchmarks or adapted to help inform specific, high-priority quality issues or initiatives within CIHI. The tool has been developed so as to allow organisations to adapt it to fit their specific data quality assessment needs. The tool contains two sets of criteria: a general set for assessing the quality of all types of data sources and a supplementary set for assessing the quality of data sources that contain survey results or sample specific data.

Within the assessment tool, the five dimensions of data quality, outlined in CIHI's Information Quality Framework, are broken down into characteristics that describe specific aspects of quality that must be understood to determine whether the data is fit for its intended use.[13] Each characteristic includes one or more criteria, framed as questions which facilitate the user to assess the quality of their data. It is the response to these criteria, or the degree to which they are met, that allows users to identify the strengths and limitations of a given data source.[34]

Within the assessment tool, the dimensions of data quality are defined as follows:

- **Relevance** — does the information meet users' current and potential future needs?
- **Accuracy and reliability —** does the information correctly and consistently describe what it was designed to measure?
- **Comparability and coherence** — is the information consistent over time and across providers and can it be easily combined with other sources?
- **Timeliness and punctuality —** is the information current and released on schedule?
- **Accessibility and clarity —** is the information and its supporting documentation easily accessed and clearly presented in a way that can be understood?

CIHI's five dimensions of data quality encompass those defined in the *Guidelines for the template for a generic national quality assurance* framework (NQAF),[12] which were developed by an expert international working group in conjunction with the United National Statistical Commission. There dimensions are considered to be international best practice and have been adopted by HIQA.

5. Statistics Canada's Quality assurance framework, 2017

5.1 Statistics Canada

The role of Statistics Canada is to provide credible and relevant statistical information to the public, to inform decision-making and to help Canadians better understand their country. Under the Statistics Act, Statistics Canada is required to collect, compile, analyse, abstract and publish statistical information relating to the commercial, industrial, financial, social, economic and general activities and condition of the people of Canada. Statistics Canada has two main objectives:

1. To provide statistical information and analysis about Canada's economic and social structure to:

   - develop and evaluate public policies and programs
   - improve public and private decision-making for the benefit of all Canadians.

2. To promote sound statistical standards and practices by:

   - using common concepts and classifications to provide better quality data
   - working with the provinces and territories to achieve greater efficiency in data collection and less duplication
   - reducing the burden on respondents through greater use of data sharing agreements (sources used include annual tax records, monthly employee payroll records and customs records)
   - improving statistical methods and systems through joint research studies and projects. [41]

5.2 Statistics Canada's Quality Assurance Framework (QAF)

The Quality Assurance Framework (QAF) [14] serves as the highest-level governance tool for quality management at Statistics Canada. This third edition was inspired by *Guidelines for the Template for a Generic National Quality Assurance Framework (NQAF)* [12] developed by the expert group of the United National Statistics Division. The QAF describes, in general terms, the measures that Statistics Canada has put in place to manage quality. The QAF also provides guidance to statistical program areas on the development and implementation of quality management strategies, to ensure data is fit for use. The QAF is used in conjunction with Statistics Canada management practices and sets out the principles that guide the work of statistics Canada. Underlying Statistics Canada's quality management strategies are eight guiding principles:

- **Quality is multi-dimensional** — Statistics Canada has identified six dimensions of statistical information to define its quality and evaluate its fitness for use: relevance, accuracy, timeliness, accessibility, coherence and interpretability.

- **Quality is relative, not absolute** — management of quality must be in conjunction with other important factors, including the data needs of users and stakeholders, costs and response burden.

- **Every employee has a role to play in assuring quality** — success in assuring quality at Statistics Canada requires the sound application of knowledge and expertise by employees at all levels within the organisation.

- **Quality must be built in at each phase of the process** — as operations at all stages can impact the quality of outputs, effective quality assurance requires measures at multiple phases of the statistical process and consideration of the impact of each phase on the process as a whole. Modelling the statistical process by dividing it into phases has proven to be an effective management tool. One such reference framework is the Generic Statistical Business Process Model, in which the principal phases are "specify needs", "design", "build", "execute" and "evaluate".[42]

- **Balancing the dimensions of quality is best achieved through a team approach** — the use of multidisciplinary teams ensures that the dimensions of quality and other important factors including cost and user needs are effectively managed.

- **Quality assurance measures must be adapted to the specific program** — at Statistics Canada, responsibility and accountability for quality assurance lie with the agency's statistical programs. Statistics Canada develops and maintains quality management strategies and tools that program areas can adapt to their individual needs.

- **Users must be informed of data quality so that they can judge whether the statistical information is appropriate for their particular use** — some dimensions of quality, such as timeliness, can be observed directly by users. However, for most other dimensions, users require objective information about data quality to evaluate fitness for use.

- **Quality assurance is a continuous practice** — Statistics Canada will ensure its methods remain at the forefront of those used by national statistical offices through embedding a culture that promotes a continuous search for new and innovative sources and methods.

The QAF consists of 12 stand-alone chapters, each defined by a quality management theme. Each chapter has a three-part approach:

- **Description —** details the concepts to be discussed in the chapter and provides definitions of relevant terms and context and background information. It highlights any issues for Statistics Canada in achieving their goal relating to that concept.

- **Assessment —** lists the objectives which are essential to the successful operationalisation of the concept in that chapter.

- **Implementation —** demonstration of the achievement of the objectives through reference to specific activities of Statistics Canada that contribute to quality data.

The 12 quality management themes are further sub-divided into three overarching themes and include:

The corporate environment:

- quality commitment
- sound implementation of statistical methods
- assurance of confidentiality, privacy and security.

The statistical programme:

- management of input data and relations with data providers
- allocation and management of resources
- management of relations with data users and stakeholders.

Statistical outputs:

- relevance
- accuracy and reliability
- timeliness and punctuality
- accessibility and clarity
- coherence and comparability
- interpretability and management of metadata.

5.2.1 Statistical outputs — dimensions of data quality

The theme of statistical outputs most closely aligns to the dimensions of data quality which HIQA uses to assess data quality. Theses dimensions are defined by Statistics Canada as follows:

- Relevance — the relevance of statistical information reflects the degree to which it meets the needs of data users and stakeholders.
- Accuracy and reliability — the accuracy of statistical information refers to the degree to which it correctly describes the phenomena it was designed to measure. Reliability reflects the degree to which statistical information, consistently over time, correctly describes the phenomena it was designed to measure.
- Timeliness and punctuality — timeliness of data is defined as the length of time between the end of the reference period or the reference date to which the data relate and the date the product is made available. Punctuality refers to the difference between the planned and actual availability.
- Accessibility and clarity — accessibility refers to the ease with which users are able to identify, obtain and use statistical products and services. Clarity refers to the degree to which metadata and other information are provided so that users are able to locate and select products or services that correspond to their needs.
- Coherence and comparability — coherence of statistics refers to the extent to which they are logically consistent in terms of definition and measurement and thus can be reliably combined in different ways and for various uses. Comparability refers to the

extent to which differences over time or among sources can be attributed to changes in the true values of the statistics, and not to changes in definition or measurement.

▪ Interpretability and management of metadata — managing interpretability primarily involves providing metadata so that statistical information can be understood and used appropriately.

Interpretability and management of metadata is not included as a separate dimension by HIQA, but this concept is covered within coherence and comparability instead.

6. Australian Capital Territory (ACT) Health, Data quality framework, 2013

6.1 ACT Health

In Australia, the states and territories have primary responsibility for public hospitals and community and public health, ambulance, public dental services and mental health programs. ACT Health, a territory in south-east Australia which contains Canberra, holds responsibility for the provision of such programs and services.

Quality data 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.

The *Health Directorate Information Management Framework 2012-2014* details the six information management principles that inform the ACT Health's approach to managing information[15]:

- **Purpose** — authoritative information is provided to meet the needs and priorities of the ACT Government and the people that use health 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.

6.2 Data quality framework components

ACT Health developed a regional data quality framework in 2013.[15] The framework draws from the ABS Data Quality Framework[16] and data quality framework developed by the CIHI[10], providing an objective approach to assessing and improving the quality of data created and managed by ACT health. 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 components of the data quality framework include[15]:

- the data integrity strategy
- the data quality policy
- the data quality improvement cycle
- the data quality assessment tool.

The data quality framework contains additional tools, including a data quality indicator tool and the ACT Health data quality statement to support improvements in the quality and management of health information.

6.2.1 The data integrity strategy

The data integrity strategy conveys the direction for data quality improvement in ACT Health. The strategy details and prioritises a number of initiatives that strengthen the directorates approach to information collection, collation, use and security.

6.2.2 The data quality policy

The data quality policy is the instrument that authorises the framework approach. The policy focuses on ensuring that staff apply rigour within their roles to accurately record, update and maintain a credible source of information over time. 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.

6.2.3 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. This ongoing approach consists of a number of phases or stages of analysing and assessing input, throughput, and output while continually moving to improve the quality.

The data quality improvement cycle's approach is to:

- **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.

6.2.4 The data quality assessment tool

The data quality assessment tool enables a comprehensive assessment of the fitness for purpose of the data within the directorate. The data quality assessment tool comprises seven dimensions of data quality, each of which is divided into a number of data characteristics that may be used to systematically assess the fitness for purpose of data. The dimensions used within the tool are adopted from the Australian Bureau of Statistics. It is intended that the tool will enable 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 to design a data collection or product which is fit for purpose. The data quality dimensions used in this framework are:

- **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 it 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. It 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.

The seven dimensions are not mutually exclusive and the importance of each dimension is not necessarily equal and may vary depending on the data and its content.[15]

7. Australian Bureau of Statistics, Data quality framework, 2009

7.1 Australian Bureau of Statistics (ABS)

The Australian Bureau of Statistics (ABS) is Australia's national statistical agency. It provides trusted official statistics on a wide range of economic, social, population and environmental matters of importance to Australia.

The ABS also has an important leadership role, coordinating statistical activities and collaborating with official bodies in the collection, compilation, analysis and distribution of statistics. This assists in maximising the value of government investment on these activities, and ensures effective and safe use of statistics which are fit-for-purpose.[43]

7.2 ABS Data Quality Framework

The ABS Data Quality Framework[16] provides the standards for assessing and reporting on the quality of statistical information. It can also assist with the development of statistical collections to produce high-quality outputs.

The ABS Data Quality Framework is based on the Statistics Canada quality assurance framework[14] and the European Statistics Code of Practice.[44] It has been designed to be used in evaluating the quality of statistical collections and products (for example, survey data and statistical tables), including administrative data. The framework is comprised of seven dimensions of quality, reflecting a broad and inclusive approach to quality definition and assessment. The seven dimensions of quality used in this framework are institutional environment, relevance, timeliness, accuracy, coherence, interpretability and accessibility. The framework provides a description for each dimension, followed by key aspects to consider when evaluating the quality of statistics against that dimension as well as suggested questions to assess quality under that dimension. All seven dimensions should be included for the purpose of quality assessment and reporting. However, the seven dimensions are not necessarily equally weighted as the importance of each dimension may vary depending on the data source and context.

The dimensions within the framework are defined as follows:

- **Institutional environment** — the institutional and organisational factors which may have a significant influence on the effectiveness and credibility of the agency producing the statistics. Consideration of the institutional environment associated with a statistical product is important as it enables an assessment of the surrounding context, which may influence the validity, reliability or appropriateness of the product.
- **Relevance** — how well the statistical product or release meets the needs of users in terms of the concept(s) measured and the population(s) represented. Consideration of the relevance associated with a statistical product is important as it enables an

assessment of whether the product addresses the issues most important to policymakers, researchers and the broader Australian community.

- **Timeliness** — the delay between the reference period (to which the data pertain) and the date at which the data become available and the delay between the advertised date and the date at which the data become available (that is, the actual release date). These aspects are important considerations in assessing quality as lengthy delays between the reference period and data availability or between advertised and actual release dates can have implications for the currency or reliability of the data.

- **Accuracy** — the degree to which the data correctly describe the phenomenon they were designed to measure. This is an important component of quality as it relates to how well the data portray reality, which has clear implications for how useful and meaningful the data will be for interpretation or further analysis. In particular, when using administrative data, it is important to remember that statistical outputs for analysis are generally not the primary reason for the collection of the data.

- **Coherence** — the internal consistency of a statistical collection, product or release, as well as its comparability with other sources of information, within a broad analytical framework and over time. The use of standard concepts, classifications and target populations promotes coherence, as does the use of common methodology across surveys. Coherence is an important component of quality as it provides an indication of whether the dataset can be usefully compared with other sources.

- **Interpretability** — the availability of information to help provide insight into the data. Information available which could assist interpretation may include the variables used, the availability of metadata, including concepts, classifications and measures of accuracy. Interpretability is an important component of quality as it enables the information to be understood and utilised appropriately.

- **Accessibility** — the ease of access to data by users, including the ease with which the existence of information can be ascertained as well as the suitability of the form or medium through which information can be accessed. The cost of the information may also represent an aspect of accessibility for some users. Accessibility is a key component of quality as it relates directly to the capacity of users to identify the availability of relevant information, and then to access it in a convenient and suitable manner.

The ABS Data Quality Framework is a general framework to enable a comprehensive and multi-dimensional assessment of the quality of a statistical dataset, product or release. It is intended that the framework enable data users and providers to:

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

The ABS recommends that where users of statistics are assessing the quality of a data item, dataset or other statistical product, using the data quality framework, a quality statement should be developed. A quality statement is a presentation of information about the quality

of a data item or a collection of data items. 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. For example, the ABS has produced specific quality statements based on statistical releases called "quality declarations". Quality declarations are succinct summaries which quickly communicate key statistical quality messages, as well as providing links to more detailed information about statistical output. ABS quality declarations are designed primarily for electronic dissemination; therefore, they are more concise in length and enable layering of information in a web environment whereby each successive layer contains more detailed information. Quality declarations complement but do not replace the more comprehensive and complete ABS quality statements.[45]

7.3 Data Quality Online Tool

The ABS, as part of a National Statistical Service (NSS) initiative, created an online assistant called the Data Quality Online Tool to help people understand data quality and management.[46] The online tool expands upon the ABS Data Quality Framework, providing users with information on the uses of the framework with the purpose of assisting them in making quality informed decisions. The data quality online tool can be found at www.nss.gov.au. It provides conceptual information in the form of questions for each of the seven dimensions of the ABS Data Quality Framework to help users to determine whether data is fit for their needs. The online tool explains how the ABS Data Quality Framework can be used to:

- define a data need
- declare quality of your own, or another organisation's data item or collection of data items (quality statements)
- compare your data need with available data (identification of data gaps)
- design a collection (this may be a result of the identification of data gaps).

8. European Statistical System Quality assurance framework

8.1 European Statistical System (ESS)

The European Statistical System (ESS) is the partnership between Eurostat and the national statistical institutes and other national authorities responsible for the development, production and dissemination of European statistics in each EU member state. This partnership also includes the European Economic Area (EEA) and the European Free Trade Association (EFTA) countries. Member states collect data and compile statistics for national and EU purposes. The ESS functions as a network in which Eurostat's role is to lead the way in the harmonisation of statistics in close cooperation with the national statistical authorities. ESS work concentrates mainly on EU policy areas but, with the extension of EU policies, harmonisation has been extended to nearly all statistical fields.

The ESS also coordinates its work with candidate countries and at European level with other services, agencies and international organisations such as OECD, the UN, the International Monetary Fund and the World Bank.[47]

8.2 Quality Assurance Framework of ESS

The Quality Assurance Framework of the ESS[11] is a supporting document aimed at assisting the implementation of the European Statistics Code of Practice[44] through identification of possible activities, methods and tools that can provide guidance and evidence for the implementation of the European Statistics Code of Practice. The European Statistics Code of Practice is based on 15 principles. However, the ESS Quality Assurance Framework covers all principles of the Code of Practice which are related to statistical processes (principles 7 to 10) and to statistical output (principles 11 to 15). Principle 4, Commitment to Quality, is also covered as it sets important fundamentals for the aforementioned principles.

The Quality Assurance Framework is divided into 3 key themes:

Institutional environment:

 ▪ commitment to quality

Statistical processes:

 ▪ sound methodology
 ▪ appropriate statistical procedures
 ▪ non-excessive burden on respondents
 ▪ cost effectiveness.

Statistical output:

 ▪ relevance
 ▪ accuracy and reliability

- timeliness and punctuality
- coherence and comparability
- accessibility and clarity.

A set of indicators demonstrating how compliance can be achieved accompany each of the principles.

8.2.1 Statistical output

The ESS outlines the following in relation to statistical outputs:

> Available statistics meet users' needs. Statistics comply with the European quality standards and serve the needs of European institutions, governments, research institutions, business concerns and the public generally. The important issues concern the extent to which the statistics are relevant, accurate and reliable, timely, coherent, comparable across regions and countries, and readily accessible by users.[11]

The principles that fall under the theme of statistical output most closely align to the dimensions of data quality used in data quality frameworks. Within the theme each principle has a number of associated indicators, with a number of methods at the institutional level and at product/survey level which must be addressed in order to satisfy that indicator.

The principles under each indicator are as follows:

Relevance — European statistics meet the needs of users:

- processes are in place to consult users, monitor the relevance and utility of existing statistics in meeting their needs, and consider their emerging needs and priorities
- priority needs are being met and reflected in the work programme
- user satisfaction is monitored on a regular basis and is systematically followed up.

Accuracy and Reliability — European statistics accurately and reliably portray reality:

- source data, intermediate results and statistical outputs are regularly assessed and validated
- sampling errors and non-sampling errors are measured and systematically documented according to the European standards
- revisions are regularly analysed in order to improve statistical processes.

Timeliness and Punctuality — European statistics are released in a timely and punctual manner:

- timeliness meets European and other international release standards
- a standard daily time for the release of European Statistics is made public
- the periodicity of statistics takes into account user requirements as much as possible

- divergence from the dissemination time schedule is publicised in advance, explained and a new release date set
- preliminary results of acceptable aggregate accuracy can be released when considered useful.

Coherence and Comparability — European statistics are consistent internally, over time and comparable between regions and countries; it is possible to combine and make joint use of related data from different sources:

- statistics are internally coherent and consistent (that is, arithmetic and accounting identities observed)
- statistics are comparable over a reasonable period of time
- statistics are compiled on the basis of common standards with respect to scope, definitions, units and classifications in the different surveys and sources
- statistics from different sources and of different periodicity are compared and reconciled
- cross-national comparability of the data is ensured within the European Statistical System through periodical exchanges between the European Statistical System and other statistical systems. Methodological studies are carried out in close co-operation between the Member States and Eurostat.

Accessibility and Clarity — European statistics are presented in a clear and understandable form, released in a suitable and convenient manner, available and accessible on an impartial basis with supporting metadata and guidance:

- statistics and the corresponding metadata are presented and archived in a form that facilitates proper interpretation and meaningful comparisons
- dissemination services use modern information and communication technology and, if appropriate, traditional hard copy
- custom-designed analyses are provided when feasible and the public is informed.
- access to microdata is allowed for research purposes and is subject to specific rules or protocols
- metadata are documented according to standardised metadata systems
- users are kept informed about the methodology of statistical processes, including the use of administrative data
- users are kept informed about the quality of statistical outputs with respect to the quality criteria for European Statistics.

8.3 Quality and performance indicators, 2014

The ESS Quality and performance indicators[48] were published in 2014, having been reviewed by the Eurostat Expert Group on Quality Indicators in 2010 and by the Task Force on Quality Reporting in 2012–2013. The standard quality and performance indicators listed in Table 5 may be used for summarising the quality of statistical products in various statistical domains.

Table 5. The European Statistical System(ESS) quality and performance indicators

Quality component	Concept	Quality and performance indicators
Relevance	Completeness	Data completeness — rate
Accuracy	Sampling error	Sampling error — indicators
	Non-sampling error	Over-coverage — rate
		Unit non-response — rate
		Item non-response — rate
		Imputation — rate
		Common units — proportion
	Revision practice	Data revision — average size
Timeliness and punctuality	Timeliness	Time lag — first results
		Time lag — final results
	Punctuality	Punctuality — delivery and publication
Accessibility and clarity	Metadata	Metadata — consultation
	On-line database	Data tables — consultations
	Documentation on quality and methodology	Metadata completeness — rate
Coherence and comparability	Comparability over time	Length of comparable time series
	Comparability – geographical (over countries, regions etc.)	Asymmetry for mirror flows statistics — coefficient

9. UN Statistics Division Template for a national quality assurance framework, 2012

9.1 United Nations Statistics Division (UNSD)

The UN Statistics Division (UNSD) is a global statistical center responsible for the coordination of international statistical activities. UNSD supports the functioning of the United Nations Statistical Commission as the apex entity of the global statistical system. UNSD has responsibility for compiling and disseminating global statistical information, developing standards and norms for statistical activities, and supporting countries' efforts to strengthen their national statistical systems.

The main functions of the UNSD are:

- **Data** — collection, processing and dissemination of statistical information
- **Methodology** — standardisation of statistical methods, classifications and definitions
- **Capacity development** — technical cooperation programme
- **Coordination** — coordination of international statistical programmes and activities.

To carry out its functions, the UNSD:

- provides a global centre for data on international trade, national accounts, energy, industry, environment and demographic and social statistics gathered from national and international sources
- promotes international standards of methods, classifications and definitions used by national agencies
- assists Member States, at their request, to improve their statistical services by giving advice and training
- coordinates international statistical programmes and activities entrusted to the Division by the United Nations Statistical Commission and the Committee for the Coordination of Statistical Activities
- provides input and secretarial support to the United Nations Statistical Commission
- facilitates the follow-up and review process of the 2030 Agenda for Sustainable Development, acting as Secretariat of the Inter-agency and Expert Group on Sustainable Development Goals (SDG) indicators and maintaining the global SDG indicators database
- promotes modern surveying and mapping techniques as a tool for growth and development.[49]

9.2 Template for a National Quality Assurance Framework (NQAF)

The development of the Template for a Generic NQAF and the Guidelines to Accompany the Template[40] was undertaken by the Expert Group on NQAF in response to a request by the United Nations Statistical Commission at its forty-first session in 2010. The template is

intended to be used as a tool to provide the general structure within which countries can formulate and operationalise national quality frameworks of their own or further enhance their existing quality frameworks. The template is divided into five areas:

- quality context
- quality concepts and frameworks
- quality assurance guidelines
- quality assessment and reporting
- quality and other management frameworks.

9.2.1 Quality context

UNSD outline the need for every organisation to have a quality management system (or equivalent) in place to ensure quality in processes and outputs. A report by Statistics Canada to the Statistical Commission in 2010[50] set the stage for the expert group to develop a generic national quality assurance framework template to provide insight into what can be included in a national quality assurance framework. The report advocated that all national statistical organisations should have a national quality assurance framework in place or consider developing one if they have not yet done so. The main benefits of having a quality assurance framework in place are:

- it provides a systematic mechanism for facilitating the ongoing identification of quality problems and possible actions for their resolution, while, at the same time, it serves to stimulate and maximise the interaction among staff throughout the organisation
- it gives greater transparency to the processes by which quality is assured and reinforces the image of the office as a credible provider of good quality statistics
- it provides a basis for creating and maintaining a quality culture within the organisation and contains reference material that can be helpful for training
- it supports quality improvements and their maintenance over time
- it is a mechanism for the exchange of ideas on quality management with other producers of statistics within the national statistical system and with other national and international statistical organisations.

9.2.2 Quality concepts and frameworks

The expert group compiled an online glossary of quality-related terms, the main source of which was the Statistical Data and Metadata eXchange (SDMX) Metadata Common Vocabulary[51] that was developed by a partnership of international organisations. Since the SDMX Metadata Common Vocabulary is an agreed global standard, preference was given to the definitions it presents in cases where several definitions were available for the same terms. In addition to the definition of each term, the NQAF Glossary includes information whenever possible in the "context" field to provide additional explanations or other useful information that contributes to the understanding of the concepts.

A summary table showing the correspondence among the NQAF, the European Statistics Code of Practice (CoP),[44] the International Monetary Fund's Data Quality Assessment Framework (DQAF),[52] Statistics Canada's Quality Assurance Framework,[14] and the Proposal for a code of good practice in statistics in Latin America and the Caribbean[53] is included as an appendix to the NQAF.

9.2.3 Quality assurance guidelines

The framework consists of 19 specific guidelines in relation to management of statistical data. The guidelines are divided into four broad themes[40]:

- managing the statistical system
- managing the institutional environment
- managing statistical processes
- managing statistical outputs.

Each guideline is accompanied by a description, "elements to be assured" at specific levels (for example, at a national statistical system level or at the agency level) and "supporting mechanisms" required in order to satisfy the guideline. The theme of managing statistical outputs closely aligns to the dimensions of data quality used by HIQA and internationally as part of a data quality framework. The guidelines within this theme include:

- assuring relevance
- assuring accuracy and reliability
- assuring timeliness and punctuality
- assuring accessibility and clarity
- assuring coherence and comparability
- managing metadata.

Each guideline is comprised of a definition and description of that dimension, elements to be assured at the agency level, the programme design stage, the programme implementation stage and at the post-collection evaluation stage. The "elements to be assured" are a list of questions to guide users in assessing if their data is relevant.

9.2.4 Quality assessment and reporting

The systematic quality assessment of a statistical agency's statistics, both from the statistical production process and product point of view, permits the agency to control and evaluate the various statistical sub-processes such as data collection, editing or weighting and thus be confident that the possible problems will be detected. The purpose is to have mechanisms in place in order to prevent, reduce and evaluate problems that may arise during the statistical process and affect the statistical products. Having solid approaches to quality assessment supports the agency's claims of being professional and credible as a producer of high-quality data.

A brief outline of what is involved in the assessment or evaluation of quality should be provided, that is, the set of information on which the quality assessment is based, the

different ways in which quality assessments can be conducted, how quality assessment can contribute to standardisation and to continuous quality improvement within a statistical agency. This includes:

- measuring product and process quality — use of indicators, quality targets and process variables and descriptions
- communicating about quality — quality reports
- obtaining feedback from users
- conducting assessments — labelling and certification
- assuring continuous quality improvement.

9.2.5 Quality and other management frameworks

The management of quality is an integral part of the management of every programme within an agency and an important component of the agency's management as a whole. Various frameworks will exist within an agency, and it is important that these frameworks are not developed and implemented in isolation, but rather, work towards achieving good quality through close interaction. The following are some of the quality and other management frameworks which are important to quality:

- performance management
- resource management
- ethical standards
- continuous improvement
- governance.

10. Quality framework and guidelines for Organisation for Economic Co-operation and Development (OECD) statistical activities, 2011

10.1 Organisation for Economic Co-operation and Development (OECD)

The OECD provides a forum in which governments can work together to share experiences and seek solutions to common problems. It works with governments to understand what drives economic, social and environmental change. It measures productivity and global flows of trade and investment. The OECD also analyses and compares data to predict future trends and sets international standards on a wide range of things, from agriculture and tax to the safety of chemicals.[54]

OECD statistics serve as a key input for analytical work that informs policy recommendations and policy-making and also as an output in their own right in the form of publications and other products. OECD statistics cover a wide range of domains, including economics, health, education, tax, agriculture, the environment, migration and quality of life.[17]

10.2 Quality framework and guidelines for statistical activities

The OECD Quality Framework and Guidelines for Statistical Activities[17] focuses on improving the quality of data collected, compiled and disseminated by the OECD through an improvement in the organisation's internal statistical processes and management. Furthermore, there will also be a positive knock-on effect on the quality of data compiled at the national level. In this sense, the OECD quality initiative is similar to those developed by Statistics Canada and other organisations, which also encompass statistical managerial and technical processes. The OECD Quality Framework has four elements:

1. a definition of quality and its dimensions
2. a procedure for assuring the quality of proposed new statistical activities
3. a procedure for evaluating the quality of existing statistical activities on a regular basis
4. a set of broad principles on which OECD statistical activities are to be conducted and quality guidelines covering all phases of the statistical production process.

The Quality Framework and Guidelines is laid out as follows:

- **Part 1** — definitions of data quality and a description of procedures for assuring the quality of new statistical activities for evaluating the quality of existing statistical activities. Part 1 also includes a set of basic principles on which OECD statistical activities must be conducted.
- **Part 2** — quality guidelines for different phases of the statistical production process.

10.2.1 Quality and its dimensions

Quality is defined as fitness for use in terms of user needs. This definition is broader than has been used in the past, when quality was equated with accuracy. It is now generally recognised that there are other important dimensions. Even if data is accurate, they cannot be said to be of good quality if they are produced too late to be useful, cannot be easily accessed or appear to conflict with other data. Thus, quality is viewed as a multi-faceted concept. The quality characteristics of most importance depend on user perspectives, needs and priorities, which vary across groups of users. Given the work already done by several statistical organisations, the OECD drew on their experience and adapted it to the organisation's context. Thus, the OECD views quality in terms of the following seven dimensions:

- relevance
- accuracy
- credibility
- timeliness
- accessibility
- interpretability
- coherence.

Another factor is that of cost-efficiency, which, while not strictly speaking a quality dimension, is still an important consideration in the possible application of one or more of the seven dimensions cited previously to OECD statistical output.

Each dimension is accompanied by a description of the dimension, considerations under that dimension and some detail of considerations within the OECD context.

10.2.2 Procedure for assuring the quality of proposed new statistical activities

The main steps in the development of a new statistical activity were defined as:

- definition of the data requirements in general terms
- evaluation of other data currently available
- planning and design of the statistical activity
- extraction of data and metadata from databases within and external to OECD
- implementation of specific data and metadata collection mechanism
- data and metadata verification, analysis and evaluation
- data and metadata dissemination.

The quality concerns and instruments available to address those concerns were identified for each step. A set of guidelines and concrete procedures have been prepared for each step which take into account good practice within the OECD and other statistical agencies.

10.2.3 Procedure for evaluating the quality of existing statistical activities on a regular basis

The procedure for reviewing the quality of existing statistical activities conducted across the OECD takes into account the fact that the review will be carried out on a rotation basis over a number of years. The stages envisaged are as follows:

- identification of the statistical activities for review during the course of the year
- self-assessment by the statistical activity manager and staff, resulting in a report that includes a brief summary of quality problems and a prioritised list of possible improvements together with an assessment of additional resources required for their implementation
- review of and comments on the self-assessment report by major users
- review of and comments on the self-assessment report by statistical, information technology and dissemination staff
- preparation of the final quality report, combining all comments
- discussion and resolution of any concerns about the report
- assignment of resources for selected quality improvement initiatives
- feedback by the chief statistician to stakeholders on the quality improvement initiatives proposed and the plans for their implementation.

10.2.4 Quality guidelines covering all phases of the statistical production process

The quality guidelines presented reflect known good practice adopted in the OECD Secretariat or in other national and international bodies active in statistics. They are designed to fill an information gap existing across OECD Directorates by helping statisticians to identify the highest quality and the most effective solutions for carrying out their activities. They are intended for internal use within the OECD and should be implemented by all parts of the organisation according to the nature of individual activities and available resources. The quality guidelines provided underpin the OECD Quality Framework by providing a basis for the evaluation of a proposed new statistical activity and for the self-assessment of an exiating statistical activity. While the quality of data obtained from national agencies is a very important aspect of the quality of OECD output, it is not the subject of these guidelines, which are primarily concerned with those activities that are completely under the OECD control.

The quality guidelines are aimed at a single statistical activity. However, statistical activities do not exist in isolation. They are bound together by a common statistical infrastructure and the fact that their outputs may be viewed and used in combination. Potential improvements to statistical infrastructure are the subject of other OECD initiatives. However, whilst the guidelines do not explicitly cover the infrastructure supporting statistical activities — computing, methodology, and publishing — they do take into account the interaction between statistical activities through consideration of the data they jointly make available.

A statistical activity is considered in terms of seven phases. The guidelines are structured under these phases and guidelines given separately for each one. The phases are:

- definition of the data requirements in general terms
- evaluation of other data currently available
- planning and design of the statistical activity
- extraction of data and metadata from databases within and external to the OECD
- implementation of a specific data and metadata collection mechanism
- data and metadata verification, analysis and evaluation
- data and metadata dissemination.

Each phase is considered within the context of the seven quality dimensions. The guidelines for each phase comprise aims, context and a list of the guidelines to be applied. Guidelines are provided for each phase separately.

11. South African Statistical Quality Assessment Framework, 2010

11.1 Statistics South Africa (Stats SA)

Statistics South Africa (Stats SA) is the national statistical service of South Africa. Its goal is the production of timely, accurate, and official statistics in order to advance economic growth, development, and democracy. To this end, Statistics SA produces official demographic, economic, and social censuses and surveys.[19]

11.2 South African Statistical Quality Assessment Framework (SASQAF)

The South African Statistical Quality Assessment Framework (SASQAF)[19] provides a framework and criteria used for evaluating and certifying statistics produced by government departments and other organs of state and, in some circumstances, by non-governmental institutions and organisations. The main purpose of the SASQAF is to provide a flexible structure for the assessment of statistical products. The SASQAF can be used for:

- self-assessment by producers of statistics
- reviews performed by a data quality assessment team
- assessment by data users, based on the producing agency's quality declaration
- assessment by international agencies, based on the quality declaration.

11.2.1 Dimensions of data quality

Within its framework Stats SA defines data quality in terms of fitness for use. Data quality is further defined in terms of prerequisites of quality and the eight dimensions of data quality, as follows:

- **Prerequisites of quality** — this refers to the institutional and organisational conditions that have impact on data quality. It defines the minimum set of necessary conditions that have to be met in order to produce good quality statistics. It, therefore, serves as a foundation on which all other dimensions of data quality should be premised on.
- **Relevance** — relevance of statistical information reflects the degree to which the data meets the real needs of clients. It is concerned with whether the available information sheds light on the issues of most importance to users.
- **Accuracy** — the accuracy of statistical information is the degree to which the output correctly describes the phenomena it was designed to measure.
- **Timeliness** — timeliness of statistical information refers to the delay between the reference point to which the information pertains and the date on which the information becomes available. Timeliness also addresses aspects of periodicity and punctuality activities within the statistical value chain.
- **Accessibility** — the accessibility of statistical information and metadata refers to the ease with which it can be obtained from the agency. This includes the ease with

which the existence of information can be ascertained as well as the suitability of the form or medium through which the information can be accessed. The cost of the information may also be an aspect of accessibility for some users.

- **Interpretability** — interpretability of statistical information refers to the ease with which users understand statistical information though the provision if metadata.
- **Comparability and coherence** — comparability of statistical information is the ability to compare statistics on the same characteristic between different points in time, geographical areas or statistical domains. The coherence of statistical information reflects the degree to which it can be successfully brought together with other similar statistical information from different sources within a broad analytic framework and over time.
- **Methodological soundness** — this refers to the application of national, international or peer-agreed standards, guidelines and practices to produce statistical outputs. Application of such standards fosters national and international comparability.
- **Integrity** — the integrity of statistical information refers to values and related practices that maintain users' confidence in the agency producing statistics and ultimately in the statistical product. This includes, among others, the need for the statistical system to be based on the United Nations (UN) Principles of Official Statistics[55] and includes principles of objectivity in collection, compilation and dissemination of data to ensure unbiased statistics which are not subject to confidentiality breaches or premature releases.

The aforementioned quality dimensions against which data quality is assessed are laid out in detail in the SASQAF, as follows:

- descriptions and key components of the dimension
- indicators relating to that dimension
- standards that must to be adhered to in relation to a quality indicator
- benchmarks for the standards, which determine the four different quality levels as prescribed.[19]

11.2.2 Certification of national statistics

The SASQAF certifies national statistics on one of four quality levels, these are:

- **Level four: Quality statistics** — these are statistics that meet all the quality requirements as set out in the SASQAF. They are designated as quality statistics to extent that deductions can be made from them and they are fit for use for the purpose for which they were designed.
- **Level three: Acceptable statistics** — these are statistics that meet most, but not all the quality requirements as stipulated in the SASQAF. They are designated as acceptable to the extent that, despite their limitations, deductions can be made and they are fit for use for the purpose for which they were designed.
- **Level two: Questionable statistics** — these are statistics that meet few of the quality requirements as stipulated in the SASQAF. They are designated as

questionable to the extent that very limited deductions can be made and they are, therefore, not fit for use for the purpose for which they were designed.

- **Level three: Poor statistics** — these are statistics that meet almost none of the quality requirements as stipulated in the SASQAF. They are designated as poor statistics to the extent that no deductions can be made from them and they are not fit for use for the purpose for which they were designed.[19]

12. Summary review of additional data quality frameworks

12.1 Independent Hospital Pricing Authority (IHPA), Data quality assurance framework, 2012

12.1.1 About the Independent Hospital Pricing Authority (Australia)

The Independent Hospital Pricing Authority (IHPA) is an independent government agency established under Commonwealth legislation in 2011 as part of the National Health Reform Agreement reached by the Council of Australian Governments (COAG). Its main functions are:

- determining the National Efficient Price for public hospital services where services are funded on an annual basis
- determining the efficient cost for healthcare services provided by public hospitals where services are block funded
- publishing this and other information in a report each year for the purposes of information decisions makers in relation to the funding of public hospitals.

This work is required to be evidence-based and depends on access to high-quality, robust data.[56]

12.1.2 IPHA's Data Quality Assurance Framework

The IHPA's Data Quality Assurance Framework details the IHPA processes to monitor and ensure best possible data quality.[21] The objectives of the framework are to:

- establish the overarching principles and provide guidance on achieving data quality assurance for IHPA
- implement a standardised approach to quality assurance arrangements for the collection of data for calculation of the national efficient price
- promote a quality assurance culture within data collection and analysis systems at the IHPA
- contribute to good corporate governance practices with regard to data management.

12.1.3 Quality principles

The IHPA applies the following principles in establishing and managing its data quality assurance systems:

- customer focus
- leadership
- involvement of people
- process approach
- system approach to data management
- continual improvement

- factual approach to decision making
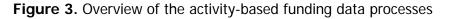- mutually beneficial supplier relationships.

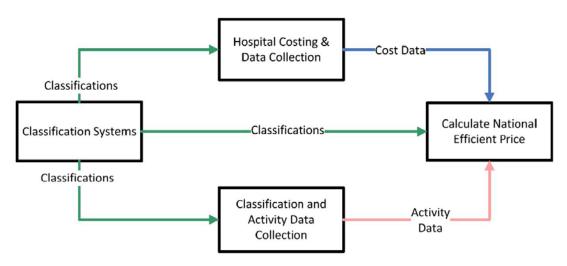## 12.1.4 Roles and responsibilities in relation to data quality assurance

The framework details the roles and responsibilities of various bodies and groups in relation to data quality assurance. The Commonwealth and the States are jointly responsible for collecting and providing data to support the objectives of comparability and transparency and to ensure that data are shared between relevant participants in national healthcare arrangements to promote better health outcomes. National Health Reform Bodies have specific requirements under the National Health Reform Agreement. These include conforming to national data development principles and balancing the national benefits of access to data against the impact on jurisdictions providing the data. The IHPA provide leadership on data quality and are responsible for functions provided under the Health Reform Act, including responsibility for ensuring data quality systems are in place to satisfy their statutory obligations.

12.1.5 Quality conformance processes

The IHPA embeds quality assurance checks into business processes to ensure that data are fit for purpose. Details of the quality assurance checks are provided for each data process that takes place as part of the activity-based funding data cycles outlined in Figure 3. The quality framework is intended to articulate ongoing quality assurance process that will deliver improved data overtime through a methodical approach to data improvement. The cycle includes processes relating to:

- activity data classification quality assurance
- activity data specification
- activity data collection and analysis
- cost data classification and specification
- cost data collection and analysis.

Figure 3. Overview of the activity-based funding data processes

Data conformance certification is prepared for all data collected for activity-based funding. This is formalised though a data quality statement for both activity and costing data.

12.2 GS1 Healthcare Data Quality Framework Version 1.0, 2017

12.2.1 About GS1

GS1 is a not-for-profit organisation dedicated to the design and implementation of global standards and solutions to improve the efficiency and visibility of supply and demand chains globally. The GS1 system of standards is the most widely used supply chain standards system in the world.

The use of GS1 standards in healthcare assists in increasing patient safety, helps drives supply chain efficiencies, provides frameworks to improve the traceability of medicines and devices, and supports clinical process.

While patient safety is paramount in healthcare, other drivers such as the ability to identify and authenticate pharmaceuticals and medical devices, track and trace products from manufacture to the point of care, improve efficiencies and enable "value chain" improvement are also vitally important.

Today, healthcare providers need to ensure high-quality patient care while working within time and financial constraints. GS1 standards and barcodes can help hospitals deliver improved patient safety and operational efficiencies.

GS1 standards help to support improved patient safety, provide data interoperability and drive business efficiency, effectiveness and cost reduction. The results in caregiver processes being supported and hospital staff able to spend more time with patients.

Regulatory agencies and jurisdictions in more than 65 countries also acknowledge GS1 standards enable critical traceability.[57]

12.2.2 Healthcare Data Quality Framework

The Healthcare Data Quality Framework[18] is intended for use by publishers/suppliers who are required to deliver quality data into the National Product Catalogue (NPC). It provides tools and support to participants across various areas of the supply chain within the healthcare sector. The framework outlines that it is the role of the supplier to remediate data errors that currently exist within the NPC and establish an internal data quality management system using guidance from the self-assessment scorecard. The framework was developed with the intention of positively impacting on cost and patient outcomes within the Australian HealthCare system. GS1 defines data quality to encompass the following pillars:

- completeness
- standards-based
- consistency
- accuracy
- time-stamped.

Within the Data Quality Framework there are five phases' outlined:

- Phase A: Action NPC Data Quality Report — commit to the Data Quality Framework. Action NPC Data Quality Report- begin data remediation
- Phase B: Conduct Data Quality Self-Assessment — manage and Complete Data Quality Self-Assessment using scorecard. Report results (internally)
- Phase C: Plan Data Quality Management system — plan Data Quality Management (DQM) based on self-assessment
- Phase D: Document SOPs and Implement Data Quality Management system processes. Report on data quality outcomes (internally)
- Phase E: Review and Embed Data Quality Management — review and adjust SOPs, embed DQMS into Policy and finalise SOPs and ensure DQMS authorities and succession is in place.

The final element of the framework is to ensure sustained data quality. This involves conducting a second Data Quality (DQ) self-assessment to ensure that internal improvements have occurred. The NPC or internal DQ reports need to be monitored and the necessary remediation needs to be actioned.

12.3 NHS Digital Data quality assurance strategy 2015–2020

12.3.1 About NHS Digital

NHS Digital is the British information and technology partner for the health and care system. NHS Digital is the national provider of information, data and IT systems for commissioners, analysts and clinicians in health and social care. It is tasked with designing, delivering and managing essential technology infrastructure, data and digital services, products and standards that healthcare professionals can utilise in their everyday duties to deliver better care.

NHS Digital aims to:

- ensure that every citizen's data is protected
- establish shared architecture and standards so everyone benefits
- implement national services to meet national and local needs
- support care organisations to get the best out of technology, data and information
- enable much better use of health and care information.[58]

In addition to developing and operating IT systems, NHS Digital collects, processes and publishes data. NHS digital supports the health and care system, helping patients make informed choices about their care while ensuring their data is kept safe.

12.3.1 Data quality assurance strategy 2015-2020

NHS Digital has developed a data quality assurance strategy[20] to support the services vision that, by 2020, data quality assurance services will provide the most consistent,

comprehensive and accessible information available about the quality of national health and social care data. The main priorities of the strategy are to:

- ensure that the impact of citizens preferences for sharing personal data are analysed and reported
- establish data quality assurance as a fundamental part of standards development
- implement data quality assurance processes for all national data services
- support organisations to continually improve the quality of their data and to get the best from our data quality assurance information
- add value to health and care data by providing access to information about its quality.

The strategy aims to transform the way NHS Digital engages on data quality assurance and the way the service works on data quality assurance by:

- radically improving engagement with partners and stakeholders
- showing those who use data quality assurance services that they are listened to and their needs are responded to
- providing tools to enable data providers and users to assess data quality
- showing that the service is flexible and dynamic at taking on new challenges
- showing responsibility for professional development
- actively supporting the wider data quality assurance agenda, including the National Information Board Framework for Action.

The strategy, as detailed in Figure 4, is guided by the values of being people focused, professional, trustworthy and innovative.

Figure 4. NHS Digital Data quality assurance strategy 2015-2020

Vision
By 2020 our data quality assurance services will provide the most consistent, comprehensive and accessible information available about the quality of national health and social care data

Priorities

Ensure that the impact of citizens' preferences for sharing personal data are analysed and reported.	Establish data quality assurance as a fundamental part of standards development.	Implement data quality assurance processes for all national data services.	Support organisations to continually improve the quality of their data and to get the best from our data quality assurance information.	Add value to health and care data by providing access to information about its quality.
• Develop evidence-based methods for measuring data consistency. • Establish a standard method for analysing and reporting the reasons for, and impact of, any data inconsistencies. • Ensure quality of identifiers supports patient-requested objections.	• Develop evidence-based data quality assurance assessment criteria. • Build and maintain a library of data quality assurance assessments for use with different types of standards. • Make sure that data quality assessments are used appropriately and the results acted upon.	• Categorise the different types of data service, e.g. secondary use, primary use, reference and public-facing. • Develop and implement a data quality assurance policy for each data service type. • Develop and implement the processes, tools and products required to support each policy. • Develop KPIs to monitor the implementation and continued use of the processes, tools and products. • Ensure national data quality assurance methods can be applied at local level.	• Set up and support data quality assurance groups. • Provide easy access to evidence-based data quality improvement information via a variety of media, including an app. • Visit organisations to provide face-to-face, hands-on support for data quality improvement.	• Build and maintain a library of data quality methods and results templates for use by the different types of data service. • Include data quality assessment methods and results in all data services' metadata. • Meet, or exceed, the Open Data Institute's Standard certificate level for all data quality information.

Transforming Delivery

We will transform the way we engage on data quality assurance	We will transform the way we work on data quality assurance
• Radically improve the way we engage with our partners and stakeholders. • Show everyone who uses our data quality assurance services that that we are listening and responding to their needs. • Provide tools to enable data providers and users to assess data quality.	• Show that we are flexible and dynamic and take on new challenges. • Show that we take responsibility for professional development. • Actively support the wider data quality assurance agenda, including the National Information Board 'Framework for Action'.

Values

People focused	Professional	Trustworthy	Innovative

References

1. *The Health Act 2007*. 2007. Available from: www.irishstatutebook.ie/eli/2007/act/23/enacted/en/pdf.

2. Department of Health and Children. *Quality and Fairness, A Health System for You. Health Strategy*. 2001. Available from: http://health.gov.ie/wp-content/uploads/2014/03/strategy-report.pdf. Accessed on: 17 April 2018.

3. Department of Health and Children. *Health Information: A National Strategy*. 2004. Available from: http://www.lenus.ie/hse/bitstream/10147/46386/1/1242.pdf. Accessed on: 17 April 2018.

4. Department of Health and Children. *Building a Culture of Patient Safety- Report of the Commission on Patient Safety and Quality Assurance*. 2008. Available from: http://health.gov.ie/wp-content/uploads/2014/03/en_patientsafety.pdf. Accessed on: 17 April 2018.

5. eHealth Ireland. *eHealth Strategy for Ireland*. 2013. Available from: http://www.ehealthireland.ie/Knowledge-Information-Plan/eHealth-Strategy-for-Ireland.pdf. Accessed on: 13 April 2018.

6. Health Service Executive. *Knowledge and Information Strategy. Delivering the Benefits of eHealth in Ireland.* 2015. Available from: http://www.ehealthireland.ie/Knowledge-Information-Plan/Knowledge-and-Information-Plan.pdf. Accessed on: 13 April 2018.

7. Eurostat. *Handbook on Data Quality Assessment Methods and Tools*. 2007. Available from: https://unstats.un.org/unsd/dnss/docs-nqaf/Eurostat-HANDBOOK%20ON%20DATA%20QUALITY%20ASSESSMENT%20METHODS%20AND%20TOOLS%20%20I.pdf. Accessed on: 2 June 2017.

8. Health Information and Quality Authority. *What you should know about data quality: a guide for health and social care staff*. 2012. Available from: https://www.hiqa.ie/sites/default/files/2017-01/Data-Quality-Guide.pdf. Accessed on: 27 March 2018.

9. Wang RY, Strong DM. Beyond accuracy: What data quality means to data consumers. Journal of management information systems. 1996;12(4):5-33.

10. Canadian Institute for Health Information. *The CIHI Data Quality Framework*. 2009. Available from: http://www.cihi.ca/cihiweb/en/downloads/data_quality_framework_2009_e.pdf. Accessed on: 1 June 2016.

11. European Statistical System. *Quality Assurance Framework of the European Statistical System*. 2015. Available from: http://ec.europa.eu/eurostat/documents/64157/4392716/qaf_2012-en.pdf/8bcff303-68da-43d9-aa7d-325a5bf7fb42. Accessed on: 1 June 2016.

12. United Nations Statistics Division. *Guidelines for the template for a generic national quality assurance framework (NQAF)*. 2012. Available from:

http://unstats.un.org/unsd/dnss/docs-nqaf/GUIDELINES%208%20Feb%202012.pdf.
Accessed on: 1 June 2017.

13. Canadian Institute for Health Information. *CIHI's Information Quality Framework.* 2017. Available from: https://secure.cihi.ca/free_products/IQF-summary-july-26-2017-en-web.pdf. Accessed on: 8 February 2018.

14. Statistics Canada. *Statistics Canada's Quality Assurance Framework.* 2017. Available from: https://www150.statcan.gc.ca/n1/en/pub/12-586-x/12-586-x2017001-eng.pdf?st=8KaUjMiO. Accessed on: 11 July 2018.

15. Australian Capital Territory (ACT) Government. *ACT Health Data Quality Framework.* 2013. Available from: http://health.act.gov.au/sites/default/files//Policy_and_Plan/Data%20Quality%20Framework.pdf. Accessed on: 7 April 2017.

16. Australian Bureau of Statistics. *The ABS Data Quality Framework.* 2009. Available from: http://www.abs.gov.au/websitedbs/D3310114.nsf/home/Quality:+The+ABS+data+quality+framework. Accessed on: 7 April 2017.

17. Organisation for Economic Co-operation and Development (OECD). *Quality framework and guidelines for OECD statistical activities.* 2011. Available from: http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=std/qfs(2011)1&doclanguage=en. Accessed on: 13 April 2017.

18. GS1 Australia. *Healthcare Data Quality Framework.* 2017. Available from: https://www.gs1au.org/WorkArea/DownloadAsset.aspx?id=2147490795. Accessed on: 11 July 2018.

19. Statistics South Africa. *South African Statistical Quality Assessment Framework.* 2010. Available from: http://www.statssa.gov.za/standardisation/SASQAF_Edition_2.pdf. Accessed on: 13 April 2017.

20. Health & Social Care Information Centre. *Data quality assurance strategy 2015-2020.* 2015. Available from: http://content.digital.nhs.uk/media/19015/Data-Quality-Assurance-Strategy-2015-2020/pdf/DQA-strategy-on-a-page.pdf. Accessed on: 27 June 2017.

21. Independent Hospital Pricing Authority (IHPA). *Data Quality Assurance Framework.* 2012. Available from: https://www.ihpa.gov.au/sites/g/files/net636/f/publications/dataqualityassuranceframework2012.pdf. Accessed on: 11 July 2018.

22. Health Information and Quality Authority. *International Review of Data Quality.* 2011. Available from: https://www.hiqa.ie/sites/default/files/2017-02/HI-International-Review-Data-Quality%20(1).pdf. Accessed on: 27 March 2018.

23. Health Information and Quality Authority. *International Review of National Health and Social Care Data Collections.* 2016. Available from: https://www.hiqa.ie/sites/default/files/2017-01/International-review-National-Health-Social-Care-Data-Collections-2016.pdf. Accessed on: 27 March 2018.

24. Health Information and Quality Authority. *Information management standards for national health and social care data collections.* 2017. Available from:

https://www.hiqa.ie/sites/default/files/2017-02/Information-management-standards-for-national-health-and-social-care-data-collections.pdf. Accessed on: 1 June 2017.

25. Health Information and Quality Authority. *Identification of Themes for Standards for National Health Information Sources in Ireland.* 2011. Available from: https://www.hiqa.ie/system/files/Identification-Theme-Standards-HI-Sources.pdf. Accessed on: 12 July 2018.

26. Health Information and Quality Authority. *The Guide to the Health Information and Quality Authority's review of information management practices in national health and social care data collections.* 2017. Available from: https://www.hiqa.ie/sites/default/files/2017-03/Guide-HIQA-review-information-management-practices-in-national-health-and-social-care-data-collections.pdf. Accessed on: 11 July 2018.

27. Central Statistics Office. *Surveys and Methodologies.* 2018. Available from: https://www.cso.ie/en/methods/. Accessed on: 01 August 2018.

28. National Cancer Registry Ireland. *Data Quality and Completeness at the Irish National Cancer Registry.* 2012. Available from: https://www.ncri.ie/publications/statistical-reports/data-quality-and-completeness-irish-national-cancer-registry. Accessed on: 02 August 2018.

29. Bray F, Parkin DM. Evaluation of data quality in the cancer registry: principles and methods. Part I: comparability, validity and timeliness. European journal of cancer. 2009;45(5):747-55.

30. Parkin DM, Bray F. Evaluation of data quality in the cancer registry: principles and methods Part II. Completeness. European journal of cancer. 2009;45(5):756-64.

31. Kerr K. *The development of a data quality framework and strategy for the New Zealand Ministry of Health.* 2000. Available from: https://pdfs.semanticscholar.org/566e/44eeb2437d33838f830e264fed9bcfb964da.pdf. Accessed on: 8 February 2017.

32. Batini C, Cappiello C, Francalanci C, Maurino A. Methodologies for data quality assessment and improvement. ACM computing surveys (CSUR). 2009;41(3):16.

33. David Loshin. *Developing a Data Quality Strategy.* 2007. Available from: http://www.b-eye-network.com/view/3816. Accessed on: 2 August 2017.

34. Canadian Institute for Health Information. *CIHI's Data Source Assessment Tool.* 2017. Available from: https://www.cihi.ca/sites/default/files/document/cihi-data-source-assessment-tool-en.pdf. Accessed on: 06 November 2017.

35. Canadian institute for Health Information (CIHI). *About CIHI.* 2016. Available from: https://www.cihi.ca/en/about-cihi. Accessed on: 7 April 2017.

36. Canadian Institute for Health Information. *CIHI celebrates data and information quality.* 2018. Available from: https://www.cihi.ca/en/cihi-celebrates-data-and-information-quality. Accessed on: 11 July 2018.

37. Canadian Institute for Health Information. *CIHI's Strategic Plan, 2016-2021.* 2017.

Available from: https://www.cihi.ca/sites/default/files/document/strategicplan2016-2021-enweb.pdf. Accessed on: 11 July 2018.

38. Canadian Institute for Health Information. *Data and Information Quality.* 2018 Available from: https://www.cihi.ca/en/data-quality. Accessed on: 11 July 2018.

39. Canadian Institute for Health Information. *CIHI's Information Quality Plan.* 2017. Available from: https://www.cihi.ca/sites/default/files/document/corporate-information-quality-plan-2016-en.pdf. Accessed on: 06 November 2017.

40. United Nations Statistics Division. *Guidelines for the template for a generic National Quality Assurance Framework (NQAF).* 2012. Available from: http://unstats.un.org/unsd/dnss/docs-nqaf/GUIDELINES%208%20Feb%202012.pdf. Accessed on: 1 June 2016.

41. Statistics Canada. *Mandate and objectives.* 2016. Available from: http://www.statcan.gc.ca/eng/about/mandate. Accessed on: 19 April 2017.

42. United Nations Economic Commission for Europe (UNECE). *General Statistical Business Process Model (GSBPM).* 2013. Available from: http://www1.unece.org/stat/platform/display/GSBPM/GSBPM+v5.0. Accessed on: 07 April 2017.

43. Australian Bureau of Statistics. *About the Australian Bureau of Statistics.* 2018. Available from: http://www.abs.gov.au/about?OpenDocument&ref=topBar. Accessed on: 11 July 2018.

44. European Statistical System Committee. *European Statistics Code of Practice.* 2011. Available from: http://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/-/KS-32-11-955. Accessed on: 1 June 2016.

45. Australian Bureau of Statistics. *Applying the ABS Data Quality Framework.* 2009. Available from: http://www.abs.gov.au/AUSSTATS/abs@.nsf/Latestproducts/1520.0Main%20Features10May%202009?opendocument&tabname=Summary&prodno=1520.0&issue=May%202009&num=&view=. Accessed on: 11 July 2018.

46. Australian Bureau of Statistics. *Data quality online.* Available from: https://www.nss.gov.au/dataquality/index.jsp. Accessed on: 12 April 2017.

47. Eurostat. *The European Statisitcal System.* Available from: http://ec.europa.eu/eurostat/web/european-statistical-system/overview?locale=fr. Accessed on: 12 April 2017.

48. Eurostat. *ESS Guidelines for the Implementation of the ESS Quality and Performance Indicators (QPI).* 2014. Available from: http://ec.europa.eu/eurostat/documents/64157/4373903/02-ESS-Quality-and-performance-Indicators-2014.pdf/5c996003-b770-4a7c-9c2f-bf733e6b1f31. Accessed on: 28 June 2017.

49. United Nations Statistics Division. *About Us.* Available from: https://unstats.un.org/home/about/. Accessed on: 13 April 2017.

50. Economic and Social Council. *Statistics Canada: national quality assurance frameworks.*

2009. Available from: https://unstats.un.org/unsd/statcom/doc10/2010-2-NQAF-E.pdf. Accessed on: 6 September 2017.

51. Statistical Data and Metadata Exchange (SDMX). *Metadata common vocabulary.* 2009. Available from: https://sdmx.org/wp-content/uploads/04_sdmx_cog_annex_4_mcv_2009.pdf. Accessed on: 20 April 2017.

52. International Monetary Fund (IMF). *Data Quality Assessment Framework.* 2012. Available from: http://dsbb.imf.org/pages/dqrs/DQAF.aspx. Accessed on: 20 April 2017.

53. Economic Commission for Latin America (ECLA) (CEPAL). *Proposal for a code of good practice in statistics in Latin America and the Caribbean and implementation plan for 2012-2013.* 2011. Available from: http://www.cepal.org/en/publications/16393-proposal-code-good-practice-statistics-latin-america-and-caribbean-and. Accessed on: 20 April 2017.

54. Organisation for Economic Co-operation and Development (OECD). *About the OECD.* Available from: http://www.oecd.org/about/. Accessed on: 13 April 2017.

55. United Nations Statistics Division. *Fundamental Principles of Official Statistics.* 2014. Available from: https://unstats.un.org/unsd/dnss/gp/fundprinciples.aspx. Accessed on: 20 April 2017.

56. Independent Hospital Pricing Authority (IHPA). *About IHPA.* Available from: https://www.ihpa.gov.au/who-we-are. Accessed on: 11 July 2018.

57. GS1 Australia. *Healthcare.* Available from: https://www.gs1au.org/for-your-industry/healthcare/. Accessed on: 11 July 2018.

58. NHS Digital. *Fit for 2020. Report from the NHS Digital Capability Review.* 2017. Available from: https://digital.nhs.uk/.../fit-for-2020-report-from-the-nhs-digital-capability-review. Accessed on: 11 July 2018.

Appendices

Appendix 1 — Glossary of abbreviations

Abbreviation	Explanation
ABS	Australia Bureau of Statistics
ACT	Australia Capital Territory
CCSA	Committee for the Coordination of Statistical Activities
CIHI	Canadian Institute of Health Information
COAG	Council of Australia Governments
CSO	Central Statistics Office
DQAF	data quality assessment framework
DQM	data quality management
EEA	European Economic Area
EFTA	European Free Trade Association
ESS	European Statistical System
EU	European Union
GP	general practitioner
GSBPM	General Statistical Business Process Model
HIQA	Health Information and Quality Authority
ICD	International Classification of Diseases
ICT	Information and Communications Technology
IHPA	Independent Hospital Pricing Authority
NCRI	National Cancer Registry of Ireland
NEP	National Efficient Price
NPC	National Product Catalogue
NPC	National Product Catalogue
NQAF	National Quality Assurance Framework
NSS	National Statistical Service
OECD	Organisation for Economic Co-operation and Development
QAF	Quality Assurance Framework
QNHS	Quarterly National Household Survey
SASQAF	South African Statistical Quality Assurance Framework
SDG	Sustainable Development Goals
SDMX	Statistical Data and Metadata Exchange
SOP	Standard Operating Procedure
UNSD	United Nations Statistics Division

Appendix 2 — Definitions of data quality dimensions used in the frameworks

Relevance	
Organisation	**Definition**
CIHI	Does the information meet users' current and potential future needs? The characteristics are: ▪ release and use ▪ value ▪ adaptability.
Statistics Canada	The relevance of statistical information reflects the degree to which it meets the needs of data users and stakeholders. Managing relevance requires ensuring that the agency's programs remain aligned with information needs as they evolve. Being aware of changing priorities and having the flexibility to respond to them are vital to ensuring continued relevance.
ACT Health	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. The characteristics are: ▪ scope and coverage ▪ reference period ▪ geographical detail ▪ main outputs/data items ▪ classifications and statistical standards ▪ type of estimates available ▪ other cautions.
Australian Bureau of Statistics	This dimension refers to how well the statistical product or release meets the needs of users in terms of the concept(s) measured, and the population(s) represented. Consideration of the relevance associated with a statistical product is important as it enables an assessment of whether the product addresses the issues most important to policy-makers, researchers and to the broader Australian community.

	Key aspects are: - scope and coverage - reference period - geographic detail - main outputs/data items - classifications and statistical standards - type of estimates available - other cautions.
European Statistical System	Available statistics meet users' needs. Statistics comply with the European quality standards and serve the needs of European institutions, governments, research institutions, business concerns and the public generally. The important issues concern the extent to which the statistics are relevant, accurate and reliable, timely, coherent, comparable across regions and countries, and readily accessible by users. Indicators are: - Processes are in place to consult users, monitor the relevance and utility of existing statistics in meeting their needs, and consider their emerging needs and priorities. - Priority needs are being met and reflected in the work programme. - User satisfaction is monitored on a regular basis and is systematically followed up.
United Nations Statistics Division	The relevance of statistical information reflects the degree to which the information meets the current and/or potential or emerging needs or requirements of clients, users, stakeholders, or the audience. Relevance, therefore, refers to whether the statistics that are needed are produced and whether those that are produced are in fact needed and useful, and shed light on the issues of most importance to users. Relevance also covers methodological soundness, particularly the extent to which the concepts, definitions and classifications correspond to user needs.
OECD	The relevance of data products is a qualitative assessment of the value contributed by these data. Value is characterised by the degree to which the data serves to address the purposes for which they are sought by users. It depends upon both the coverage of the required topics and the use of appropriate concepts. Value is further characterised by the merit of users' purposes in terms of the OECD mandate, the agreements with member countries and the opportunity costs of producing the data.
Statistics South Africa	Relevance of statistical information is the degree to which the data meets the real needs of the clients. It is concerned with whether the available information sheds light on the issues of importance to the users.

Accuracy and reliability	
Organisation	**Definition**
CIHI	Does the information correctly and consistently describe what it was designed to measure? The characteristics are: • coverage • item availability • lineage • capture and collection.
Statistics Canada	The accuracy of statistical estimates is usually quantified by the evaluation of different sources of error, where the magnitude of an error represents the degree of difference between the estimate and the true value. Common sources of error include coverage, non-response, measurement and processing. For estimates derived from survey data, an additional error source is sampling error, which reflects the fact that the estimates are computed from samples, rather than the entire population. Related to accuracy is the concept of reliability, which reflects the degree to which statistical information, consistently over time, correctly describes the phenomena it was designed to measure. That is, reliability characterizes repeated observations of accuracy, over time. Reliability applies both to multiple measurements of the same phenomenon (for example, preliminary, final and revised estimates) and to a series of measurements (for example, monthly estimates of the employment rate).
ACT Health	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). The characteristics are: • coverage error • sample error • non-response error • response error • other sources of errors

	▪ revisions to data. This framework does not include reliability as a dimension of data quality.
Australian Bureau of Statistics	Accuracy refers to the degree to which the data correctly describe the phenomenon they were designed to measure. This is an important component of quality as it relates to how well the data portray reality, which has clear implications for how useful and meaningful the data will be for interpretation or further analysis. Accuracy should be assessed in terms of the major sources of errors that potentially cause inaccuracy. Any factors which could impact on the validity of the information for users should be described in quality statements. Key aspects: ▪ coverage error ▪ sample error ▪ non-response error ▪ response error ▪ other sources of errors ▪ revisions to data. This framework does not include reliability as a dimension of data quality.
European Statistical System	European Statistics accurately and reliably portray reality. The indicators are: ▪ Source data, intermediate results and statistical outputs are regularly assessed and validated. ▪ Sampling errors and non-sampling errors are measured and systematically documented according to the European standards. ▪ Revisions are regularly analysed in order to improve statistical processes.
United Nations Statistics Division	Statistical agencies should develop, produce and disseminate statistics that accurately and reliably portray reality. The accuracy of statistical information reflects the degree to which the information correctly describes the phenomena it was designed to measure, that is, the degree of closeness of estimates to true values. It is usually characterized in terms of estimation of sampling and non-sampling errors. These errors are traditionally decomposed into bias (systematic error) and variance (random error) components, and reflect the major sources of error (for example, errors linked to sampling, coverage, measurement, nonresponse and processing). Reliability concerns whether the statistics consistently over time measure the reality that they are designed to represent.
OECD	The accuracy of data products is the degree to which the data correctly estimate or describe the quantities or characteristics they are designed to measure. Accuracy refers to the closeness between the values provided

	and the (unknown) true values. Accuracy has many attributes, and in practical terms there is no single aggregate or overall measure of it. Of necessity these attributes are typically measured or described in terms of the error, or the potential significance of error, introduced through individual sources.
GS1 Australia	Data values are right, at the right time. This framework does not include reliability as a dimension of data quality.
Statistics South Africa	The accuracy of statistical information is the degree to which the output correctly describes the phenomena it was designed to measure. This framework does not include reliability as a dimension of data quality.

Timeliness and punctuality	
Organisation	**Definition**
CIHI	Is the information current and released on schedule? The characteristics: ▪ release timeliness ▪ release punctuality ▪ submission timeliness ▪ processing timeliness.
Statistics Canada	The timeliness of a data product is defined as the length of time between the end of the reference period (or the reference date) to which the data relate and the date the product is made available. Punctuality refers to the difference between planned and actual availability.
ACT Health	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. The characteristics are: ▪ timing ▪ frequency. This framework does not include punctuality as a dimension of data quality.
Australian Bureau of Statistics	Timeliness refers to the delay between the reference period (to which the data pertain) and the date at which the data become available; and the delay between the advertised date and the date at which the data become available (that is, the actual release date). These aspects are important considerations in assessing quality as lengthy delays between the reference period and data availability or between advertised and actual release dates can have implications for the currency or reliability of the data. Key aspects are: ▪ timing ▪ frequency of survey. This framework does not include punctuality as a dimension of data quality.

European Statistical System	European Statistics are released in a timely and punctual manner. The indicators are: - Timeliness meets European and other international release standards. - A standard daily time for the release of European Statistics is made public. - The periodicity of statistics takes into account user requirements as much as possible. - Divergence from the dissemination time schedule is publicised in advance, explained and a new release date set. - Preliminary results of acceptable aggregate accuracy can be released when considered useful.
United Nations Statistics Division	Statistical agencies should minimize the delays in making data available. Timeliness refers to how fast — after the reference date or the end of the reference period — the data are released or made available, whether for dissemination or for further processing. Punctuality refers to whether data are delivered on the dates promised, advertised or announced (for example, in an official release calendar).
OECD	The timeliness of data products reflects the length of time between their availability and the event or phenomenon they describe, but considered in the context of the time period that permits the information to be of value and still acted upon. The concept applies equally to short-term or structural data; the only difference is the timeframe. Closely related to the dimension of timeliness, the punctuality of data products is also very important, both for national and international data providers. Punctuality implies the existence of a publication schedule and reflects the degree to which data are released in accordance with it.
GS1 Australia	Validity timeframe of data is clear. This framework does not include punctuality as a dimension of data quality.
Statistics South Africa	Timeliness of statistical information refers to the delay between the reference point to which information pertains and the date on which the information becomes available. Timeliness also addresses aspects of periodicity and punctuality activities within the statistical value chain. This framework does not include punctuality as a dimension of data quality.

Coherence and comparability	
Organisation	**Definition**

CIHI	Is the information consistent over time and across providers and can it easily be combined with other sources?
	
	The characteristics are:
	
	  ▪ concordance
	  ▪ standardisation
	  ▪ linkage
	  ▪ jurisdictional comparability
	  ▪ historical comparability.

Statistics Canada	Coherence of statistics refers to the extent to which they are logically consistent in terms of definition and measurement and thus can be reliably combined in different ways and for various uses. Comparability refers to the extent to which differences over time or among sources can be attributed to changes in the true values of the statistics and not to changes in definition or measurement.

ACT Health	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.
	
	The characteristics are:
	
	  ▪ changes to data items
	  ▪ comparison across data items
	  ▪ comparison with previous releases
	  ▪ comparison with other products available.
	
	This framework does not include comparability as a dimension of data quality.

Australian Bureau of Statistics	Coherence refers to the internal consistency of a statistical collection, product or release, as well as its comparability with other sources of information, within a broad analytical framework and over time. The use of standard concepts, classifications and target populations promotes coherence, as does the use of common methodology across surveys. Coherence is an important component of quality as it provides an indication of whether the dataset can be usefully compared with other sources to enable data compilation and comparison. It is important to note that coherence does not necessarily imply full numerical consistency, rather consistency in methods and collection standards. Quality statements of statistical measures must include a discussion of any factors

	which would affect the comparability of the data over time. Key aspects are: - changes to data items - comparison across data items - comparison with previous releases - comparison with other products available. This framework does not include comparability as a dimension of data quality.
European Statistical System	European Statistics are consistent internally, over time and comparable between regions and countries; it is possible to combine and make joint use of related data from different sources. The indicators: - Statistics are internally coherent and consistent (that is, arithmetic and accounting identities observed). - Statistics are comparable over a reasonable period of time. - Statistics are compiled on the basis of common standards with respect to scope, definitions, units and classifications in the different surveys and sources. - Statistics from different sources and of different periodicity are compared and reconciled. - Cross-national comparability of the data is ensured within the European Statistical System through periodical exchanges between the European Statistical System and other statistical systems. Methodological studies are carried out in close co-operation between the Member States and Eurostat.
Unite Nations Statistics Division	Statistical agencies should develop, produce and disseminate statistics that are consistent internally and comparable over time and are produced using common standards with respect to scope, definitions, classifications and units. It should be possible to combine and make joint use of related data from different sources.
OECD	The coherence of data products reflects the degree to which they are logically connected and mutually consistent. Coherence implies that the same term should not be used without explanation for different concepts or data items; that different terms should not be used without explanation for the same concept or data item; and that variations in methodology that might affect data values should not be made without explanation. Coherence in its loosest sense implies the data are "at least reconcilable."

	This framework does not include comparability as a dimension of data quality.
Statistics South Africa	Comparability of statistical information is the ability to compare statistics on the same characteristic between different points in time, geographical areas or statistical domains. The coherence of statistical information reflects the degree to which it can be successfully brought together with other similar statistical information from different sources within a broad analytical framework and over time. It is the extent to which differences between two sets of statistics are attributable to differences between the estimates and the true value of the statistics.

Accessibility and clarity	
Organisation	**Definition**
CIHI	Is the information and its supporting documentation easily accessed and clearly presented in a way that can be understood? The characteristics are: ▪ accessibility ▪ interpretability.
Statistics Canada	Accessibility refers to the ease with which users are able to identify, obtain and use statistical products and services. Clarity refers to the degree to which metadata and other information are provided so that users are able to locate and select products or services that correspond to their needs.
ACT Health	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. The characteristics are: ▪ accessibility to the public ▪ data products available. This framework does not include clarity as a dimension of data quality.
Australian Bureau of Statistics	Accessibility refers to the ease of access to data by users, including the ease with which the existence of information can be ascertained, as well as the suitability of the form or medium through which information can be accessed. The cost of the information may also represent an aspect of accessibility for some users. Accessibility is a key component of quality as it relates directly to the capacity of users to identify the availability of relevant information, and then to access it in a convenient and suitable manner. Key aspects: ▪ accessibility to the public ▪ data products available. This framework does not include clarity as a dimension of data quality.
European Statistical System	European Statistics are presented in a clear and understandable form, released in a suitable and convenient manner, available and accessible on an impartial basis with supporting metadata and guidance.

	The indicators are: ▪ Statistics and the corresponding metadata are presented and archived in a form that facilitates proper interpretation and meaningful comparisons. ▪ Dissemination services use modern information and communication technology and, if appropriate, traditional hard copy. ▪ Custom-designed analyses are provided when feasible and the public is informed. ▪ Access to microdata is allowed for research purposes and is subject to specific rules or protocols. ▪ Metadata are documented according to standardised metadata systems. ▪ Users are kept informed about the methodology of statistical processes including the use of administrative data. ▪ Users are kept informed about the quality of statistical outputs with respect to the quality criteria for European Statistics.
United Nations Statistics Division	Statistical agencies should ensure that the statistics and metadata they develop, produce and disseminate can be found or obtained without difficulty, are presented clearly and in such a way that they can be understood, are available and accessible to all users on an impartial and equal basis in various convenient formats, and are affordable, if not offered free of charge.
OECD	The accessibility of data products reflects how readily the data can be located and accessed from within OECD data holdings. The range of different users leads to such considerations as multiple dissemination formats and selective presentation of metadata. Thus, accessibility includes the suitability of the form in which the data are available, the media of dissemination, and the availability of metadata and user support services. It also includes the affordability of the data to users in relation to its value to them and whether the user has reasonable opportunity to know that the data are available and how to access them. This framework does not include clarity as a dimension of data quality.
Statistics South Africa	The accessibility of statistical information and metadata refers to the ease with which it can be obtained from the agency. This includes the ease with which the existence of information can be ascertained, as well as the suitability of the form or medium through which the information can be accessed. The cost of the information may also be an aspect of accessibility for some users. This framework does not include clarity as a dimension of data quality.

Health Information and Quality Authority