# Data Quality Framework

## Development Summary

October 2025

IHACPA

**IHACPA Data Quality Framework Development Summary — October 2025**

# Table of Contents

**Abbreviations and definitions** ..... **4**

**1. Introduction** ..... **6**

**2. Summary of recommendations** ..... **7**

**3. Project methodology** ..... **10**

**4. Project findings** ..... **14**

**5. Recommendations** ..... **25**

**6. Framework review methodology** ..... **28**

**Appendix A: Details of stakeholder engagement** ..... **31**

**Appendix B: Data quality objectives and requirements** ..... **34**

**Appendix C: National data collection and management principles and standards** ..... **40**

**Appendix D: Mapping of policy and guidance documents to data quality dimensions** ..... **55**

**Appendix E: List of public documents** ..... **58**

# Abbreviations and definitions

| Term | Definition |
|---|---|
| **ABF** | Activity based funding |
| **ABS** | Australian Bureau of Statistics |
| **AHPCS** | Australian Hospital Patient Costing Standards |
| **AI** | Artificial Intelligence |
| **AIHW** | Australian Institute of Health and Welfare |
| **APPs** | Australian Privacy Principles |
| **CEO** | Chief Executive Office |
| **CIHI** | Canadian Institute for Health Information |
| **Data lifecycle** | Data specification, acquisition, management and assurance |
| **Data Providers** | Providers of data in the data lifecycle |
| **DATA Scheme** | Data Availability and Transparency Act Scheme |
| **Data Users** | Users of data in the data lifecycle |
| **DRS** | Data Request Specifications |
| **DTA** | Digital Transformation Agency |
| **EQAO** | Education Quality and Accountability Office |
| **ESS** | European Statistical Systems |
| **ETL** | Extract, Transform and Load |
| **IHACPA** | Independent Health and Aged Care Pricing Authority |
| **In-scope data** | Data within IHACPA's current portfolio of work |
| **JAC** | Jurisdictional Advisory Committee |
| **Jurisdictions** | States, territories and the Commonwealth Governments |
| **NAC** | NHCDC Advisory Committee |
| **NEC** | National efficient cost |
| **NEP** | National efficient price |

| Term | Definition |
|---|---|
| **NHCDC** | National Hospital Cost Data Collection |
| **NHFB** | National Health Funding Body |
| **NHRA** | National Health Reform Agreement |
| **NHS** | National Health Service |
| **NQAF** | National Quality Assurance Framework |
| **NWAU** | National weighted activity unit |
| **OAIC** | Office of the Australian Information Commissioner |
| **ONDC** | Office of the National Data Commissioner |
| **PGPA** | Public Governance, Performance and Accountability Act 2013 |
| **QA** | Quality Assurance |
| **SDMS** | Secure Data Management System |
| **TAC** | Technical Advisory Committee |
| **TEHDAS** | Towards European Health Data Space |
| **The NHR Act** | *National Health Reform Act 2011(Cth)* |
| **The addendum** | Addendum to the National Health Reform Agreement 2020-26[1] |
| **The Bill** | Aged Care and Other Legislation Amendment (Royal Commission Response) Bill 2022 |
| **The Pricing Authority** | Governing body of the Independent Health and Aged Care Pricing Authority |

[1] The Addendum to the National Health Reform Agreement (NHRA) 2020–25 was extended for 12 months to enable the continued negotiation of the Addendum to the NHRA 2025–30.

# 1. Introduction

## 1.1 Context and background

In May 2012, the Independent Health and Aged Care Pricing Authority (IHACPA, or 'the agency') published a data quality assurance framework to monitor and ensure best possible data quality and integrity. The purpose of the framework was to establish data quality principles, implement uniform quality assurance for data collection, promote a quality assurance culture, and support good corporate governance in data management. Since 2012, IHACPA's data infrastructure and processes for data storage, collection, analysis and quality assurance have matured significantly.

Additionally, a key recommendation of the National Hospital Cost Data Collection (NHCDC) Public Sector Review 2021-22 Report was that IHACPA develop an NHCDC Data Quality Framework to improve the cost and activity data collections in consultation with the states and territories.

In response, IHACPA developed a Data Quality Framework (the framework) and an accompanying Development Summary (the summary, or this summary). The purpose of the framework is to enable the quality of data within IHACPA's portfolio of work to be assessed, understood, communicated and managed consistently. It provides systematic and methodological rigour to IHACPA's data quality processes at each stage of data specification, acquisition, management and assurance, known as the data lifecycle.

The framework is structured into key elements of: data quality principles; data quality definition; roles and responsibilities; quality assurance processes; enablers, systems and tools; and framework governance.

## 1.2 Purpose

This summary has been developed alongside IHACPA's framework to assist IHACPA in understanding, implementing and maintaining the framework in the organisation and its processes.

The purpose of this summary is to outline the process used to develop the framework, detail the findings of a targeted review into how IHACPA is currently applying elements of the framework to their data and quality assurance processes, and discuss areas in which IHACPA could improve their processes to enhance alignment with the framework and enhance data quality.

The summary documents:

- the approach used and research undertaken to develop the framework and conduct the targeted review (Section 3)
- the outcomes and findings of the targeted review into how IHACPA is currently applying key elements of the framework into its data and quality assurance processes (Section 4)
- a set of recommendations and improvement considerations (Section 5), and
- a high-level approach to refreshing the framework (Section 6).

The summary is consistent with version 1 (July 2025) of the framework.

# 2. Summary of recommendations

The purpose of the targeted review was to test IHACPA's current data processes against the framework, and to identify potential gaps that could be addressed to enhance data quality.

The review identified that, in general, the key elements of the framework appear to be well considered in IHACPA's current data processes for public hospital data. However, there are opportunities for IHACPA to improve alignment with the framework.

A summary of the gaps identified during the review and the associated proposed recommendations is provided in Table 1. The gaps reflect areas in which IHACPA is currently not aligning to the data quality principles or data quality dimensions outlined in the framework, whilst the recommendations outline potential ways IHACPA could address the gaps and improve alignment with the framework. The full list of findings and gaps (including more context and nuance) are in Section 4 of this summary, whilst the detailed recommendations are in Section 5.

### Table 1: Summary of identified gaps and proposed recommendations

| **Recommendation 1:**<br>**IHACPA to consider options to gain greater confidence on the accuracy of data provided by jurisdictions for the public hospital data submissions, either through an independent review or by requesting greater levels of information from jurisdictions supporting their conclusions.** | |
|---|---|
| **Identified gap(s)** | **Reference(s)** |
| IHACPA relies on the Statement of Assurance and Data Quality Statement to assess accuracy of data submitted by jurisdictions. However, the varied level of detail included in these documents, and IHACPA's limited visibility over the underlying processes of jurisdictions, pose challenges for IHACPA to consistently conclude on data quality, in line with their requirements. | Principle Gap 6<br>Relates to the "transparent" principle at the "data requirements" stage of the data lifecycle.<br><br>Dimensions Gap 4<br>Relates to the "accuracy" dimension at the "documentation and actions taken" step of the control cycle. |
| **Recommendation 2:**<br>**IHACPA to increase the completeness, coverage and holistic nature of documentation on internal data quality processes.** | |
| **Identified gap(s)** | **Reference(s)** |
| The consolidated national dataset is a single source of truth. IHACPA teams access this dataset, apply their own processes and data transformations for different use cases. This is to obtain dataset(s) that are fit for purpose for each key use case. It is important to clearly and | Principles Gap 1 |

| | |
|---|---|
| transparently document all adjustments made so that data limitations are well understood, and the appropriate dataset can be identified and applied to respond with confidence to ad hoc queries. | Relates to the “trusted” principle at the “data use” stage of the data lifecycle. |
| No holistic documentation covering the end-to-end data lifecycle spanning across teams was identified in the review. | Principles Gap 5<br>Relates to the “transparent” principle at the “data requirements” stage of the data lifecycle. |
| IHACPA maintains a communications document to record interactions with external stakeholders regarding data quality and limitations. However, this document does not extend to recording internal interactions within IHACPA teams regarding data quality and limitations. | Principles Gap 7<br>Relates to the “transparent” principle at the “data use” stage of the data lifecycle |
| Documentation does not always outline the rationale of business rules and checks performed, or how these business rules should be updated over time for their continued use in subsequent data collections. | Principles Gap 8<br>Relates to the “transparent” principle at the “data use” stage of the data lifecycle. |
| **Recommendation 3:**<br>**IHACPA to review policy and guidance documents and revise (where required) to include explicit definitions and goals, including measurable and actionable assessment criteria, for each data quality dimension.** | |
| **Identified gap(s)** | **Reference(s)** |
| The extent to which the data quality dimensions are defined in IHACPA’s current policy and guidance documents is variable. Several dimensions are explicitly defined, some are implicitly defined, and for some no clear definition was identified. | Dimensions Gap 1<br>Relates to the multiple dimensions at the “policy and guidance definition” step of the control cycle. |
| The description of data quality goals and explicit identification of assessment criteria is variable and, in some cases, missing. | Dimensions Gap 2<br>Relates to the multiple dimensions at the “data quality goals and assessment criteria” step of the control cycle. |

| **Recommendation 4:**<br>**IHACPA to investigate ways to improve the quality, timeliness and consistency of feedback and responses provided by data providers to queries on data quality. Options may include promoting the importance of feedback loops, formalising processes through existing committees, or other collaboration with data providers to work towards better practices on data quality processes and data integrity, as required under clause A154 of the National Health Reform Agreement (NHRA).** | |
|---|---|
| **Identified gap(s)** | **Reference(s)** |
| Stakeholder consultations identified some variation in the timing and completeness of data providers responses on issues relating to their data submission. | Principles Gap 2<br>Relates to the "efficient" principle at the "data preparation" stage of the data lifecycle. |
| **Recommendation 5:**<br>**IHACPA to clearly articulate its vision for innovation, integrating ongoing improvements in data quality with broader strategic activities in both pricing and non-pricing goals.** | |
| **Identified gap(s)** | **Reference(s)** |
| Policy and guidance documents reference the importance of innovation, but do not communicate its purpose of vision, nor include areas where IHACPA anticipates disruptions and how it will manage change. | Principles Gap 3<br>Relates to the "innovative" principle at the "data requirements" stage of the data lifecycle. |
| IHACPA provides improvements in data quality management to a number of disruptions across the health landscape, however these activities were not found to be clearly captured under IHACPA's innovation goals. | Principles Gap 4<br>Relates to the "innovative" principle at the "data requirements" stage of the data lifecycle. |
| **Recommendation 6:**<br>**IHACPA to review the framework on a regular basis.** | |
| **Identified gap(s)** | **Reference(s)** |
| Disruptions to technology, legislative or policy changes, and changes to IHACPA's strategic focus or operating environment may have implications for IHACPA's data quality governance over time. | N/A |

# 3. Project methodology

## 3.1 Project scope

The scope of the project was to:

- develop IHACPA's framework to support the improvement in data quality and integrity for all data assets within IHACPA's portfolio of work
- perform a targeted review to assess alignment of IHACPA's current data processes with the framework, and
- develop a summary outlining the process used to develop the framework, the findings of the targeted review, and recommendations for how IHACPA could improve processes to enhance data quality.

The framework, targeted review and associated recommendations were developed based on publicly available information, internal IHACPA documents provided, and through consultations with external stakeholders (via the NHCDC Advisory Committee and Technical Advisory Committee) and internal stakeholders from branches within IHACPA, including the:

- Infrastructure team
- Data Acquisition team
- Analytics team
- Pricing Implementation team
- Aged Care Pricing team
- Hospital Costing team
- Aged Care Costing team
- Classifications team, and
- Hospital Policy team.

The framework has been designed to be applicable to all data sources and functions within IHACPA's portfolio of work.

The targeted review was a high-level review of the information made available for the project. Limited datasets and codebases were made available as part of this project and therefore these were not considered as part of this project nor used to validate the accuracy of information gathered through the document review and stakeholder consultations. This review was not performed in accordance with generally accepted auditing, review, or other assurance standards in Australia and accordingly does not express any form of assurance.

As highlighted by the NHRA and other policy documents, data practices in collection, processing, and use are constantly being shaped by its users. Therefore, whilst the findings and recommendations identified as part of this project are relevant for IHACPA at the time of writing, they may be expected to change over time and therefore should be reviewed on an ongoing basis.

## 3.2 Project approach

The project approach consisted of four key components, involving agreeing the project purpose, scope and goals, collection and documentation of IHACPA's processes, the development of the framework and a targeted review into how the framework is applied across current IHACPA data processes.

**Figure 1: Project approach**


**Agreed project purpose, scope and goals**

As part of the project initiation, a kick off meeting was conducted with IHACPA to align on the project's purpose, scope, and goals, as well as to provide initial information requests for relevant documents and discuss the stakeholder engagement plan.

**Understanding and documenting IHACPA's processes**

Initial research was conducted to understand IHACPA's organisational objectives, current quality assurance processes, and potential future data needs. As the data processes relating to public hospital data are currently at a higher level of maturity than private hospital and aged care data, this research was focused on IHACPA's public hospital data processes (relating to the ABF and NHCDC data collections). Further, the focus was to understand the end-to-end data processes rather than individual detailed process maps.

The initial research involved a scan of publicly available information, internal documents provided by IHACPA within the Secure Data Management System (SDMS). Stakeholder consultations were conducted with team members within IHACPA, and with jurisdictions via the Technical Advisory Committee (TAC) and NHCDC Advisory Committee (NAC). For more information on the public documents accessed see Appendix E: List of public documents, and for more information on the stakeholder consultations, see Appendix A: Details of stakeholder engagement.

In the initial research scan, a number of process diagrams were provided which outlined team-specific actions and data processes within IHACPA. However, no holistic diagram was available that outlined the end-to-end processes spanning across multiple teams. Therefore, a process flow diagram was created to illustrate a simplified representation of IHACPA's data lifecycle and data flows between separate branches.

**Development of the framework**

Using a combination of research, stakeholder consultation and document review to develop an understanding of both the current state and desired future state of data quality, the framework was systematically developed, incorporating the following inputs:

- IHACPA's organisational objectives and vision
- current state data quality processes and issues

- relevant data collection and management principles, frameworks and standards for the Australian Government
- international examples and better practice data collection and management
- analysis of current state processes and documentation, and
- planned or potential changes to legislation, policy or process that could impact IHACPA's data quality activities.

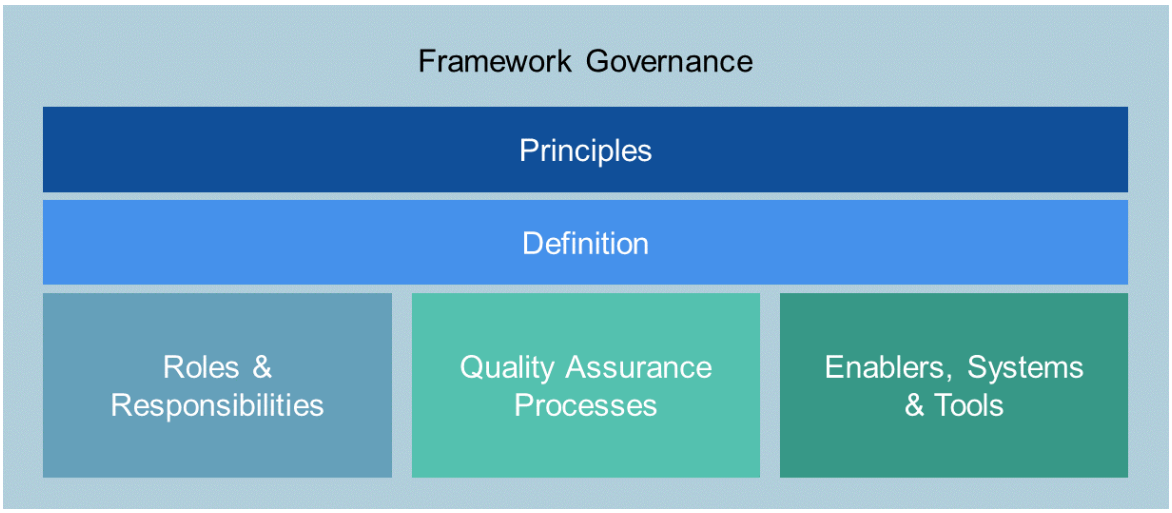
The framework has been refined in consultation with relevant internal and external stakeholders, including via NAC and TAC. For more information on IHACPA's current and future state data quality objectives, see Appendix B: Data quality objectives and requirements. For information on the relevant data collection and management principles, frameworks and standards, see Appendix C: National data collection and management principles and standards.

**Targeted review**

A targeted review was performed during and after the development of the framework to test IHACPA's current data processes against the framework, and to identify potential gaps that could be addressed to enhance data quality.

The six elements of the framework structure (Figure 2) were considered as part of the targeted review:

**Figure 2: Framework structure**




*Data Quality Principles* and *Data Quality Definition* (via the data quality dimensions):

These are foundational elements which underpin and guide the entire data quality process and are concepts that have been newly articulated in the framework. The principles reflect the various trade-offs that need to be considered when seeking to manage data quality. The Data Quality Definition (and dimensions that comprise the data quality definition) reflect the components that describe features of data quality to be achieved.

The targeted review involved mapping these foundational elements – the data quality principles and data quality dimensions – to IHACPA's current data processes. Specifically, each data quality principle was mapped to the corresponding current data process(es) at each of the six stages in the

data lifecycle. This provided an understanding of how these principles are being applied across the end-to-end data lifecycle, and identified any key gaps that exist within current processes.

For the data quality dimensions, IHACPA's current policy and guidance documents were examined to identify how each dimension is defined in these documents, whether these definitions enable measurable actions, and how IHACPA currently performs checking processes in line with the definition of each dimension. The identified gaps are outlined in Section 4 of this summary.

*Roles and Responsibilities, Quality Assurance Processes, Enablers, Systems & Tools*

These are the enablers that reflect how IHACPA carries out its data processes. These elements were mapped to the process diagram to describe where and how they influence or contribute to IHACPA's current data processes. The process diagram is discussed in Section 4 of this summary.

The review of the associated documentation and stakeholder engagement established further findings in the application of the principles and data quality definitions.

*Framework Governance*

This item relates to the development and maintenance of the framework document. Details of the high-level methodology to review the framework on an ongoing basis is outlined in Section 6 of this summary.

# 4. Project findings

This section outlines the findings identified during the targeted review of the extent to which the key elements of the framework are applied in IHACPA's current data processes for public hospital data.

The findings are broken down into the following subsections:

- **Process flow diagrams**: A summary of the end-to-end data processes within IHACPA across the data lifecycle.
- **Data quality principles**: Findings on how IHACPA is applying the framework principles in their current processes, outlining any gaps identified for each principle at each stage of the data lifecycle.
- **Data quality dimensions**: Findings on IHACPA's current design and implementation of each data quality dimension within the data quality definition stated in the framework, outlining any gaps identified for each dimension.

## 4.1 Process flow diagram

Process flow diagrams were developed in response to an identified gap in holistic documentation of end-to-end data processes within IHACPA, spanning across multiple teams at each stage of the data lifecycle. The high-level stages of the data lifecycle covered by the diagrams are shown in Table 2.

**Table 2: High-level stages of the data lifecycle**

| Stage | Description |
|---|---|
| Stage 1: Data requirements | Identifying policy and guidance documents, as well as stakeholders involved in defining IHACPA's data requirements. |
| Stage 2: Data collection | Outlining the submission process for data providers. |
| Stage 3: Data ingestion | Outlining the initial checks performed on the data. |
| Stage 4: Data validation | Outlining a feedback mechanism between IHACPA and data providers. |
| Stage 5: Data preparation | Outlining the steps required to create the national dataset. |
| Stage 6: Data use | Identifying uses of data across multiple branches within IHACPA. |

## 4.2 Data quality principles

The data quality principles are the foundation for IHACPA and its data provider's approach to data quality. They provide direction on decisions and activities relating to data quality based on IHACPA's strategic objectives – as defined in the Work Program and Corporate Plan. The principles are to be used to guide the application and maintenance of the framework, outlining the expected practices, procedures and attitudes to promote appropriate data quality maturity for data providers and IHACPA. The five principles are trusted, attainable, efficient, innovative and transparent.

This section provides a summary of findings on how IHACPA is applying these framework principles in their current processes for public hospital data.

Table 3 provides a mapping between the gaps identified, their respective stages in the data lifecycle, and the data quality principles to which they relate.

### Table 3: Summary of framework principles gaps identified

| Data lifecycle stage | Principles | | | | |
|---|---|---|---|---|---|
| | Trusted | Attainable | Efficient | Innovative | Transparent |
| Stage 1: Data requirements | - | - | - | Gap 3 and Gap 4 | Gap 5 |
| Stage 2: Data collection | - | - | - | - | Gap 6 |
| Stage 3: Data ingestion | - | - | - | - | - |
| Stage 4: Data validation | - | - | - | - | - |
| Stage 5: Data preparation | - | - | Gap 2 | - | - |
| Stage 6: Data use | Gap 1 | - | - | - | Gap 7 and Gap 8 |

"-" indicates that no significant gaps were identified during the targeted review – however this does not mean there are not potential improvements in those areas.

**Principle 1: Trusted**

> **Framework definition:** Data that is clearly defined and well-understood is collected, analysed and managed effectively, providing reliable and defensible evidence-based outputs.

IHACPA has a range of internal policy documents which support the secure and safe collection of data and analysis. For public hospital datasets, jurisdictions are also provided with tools including the data submission portal and the NHCDC dashboard, which maintain data integrity by providing accountability of the data quality at points in time.

*Principles Gap 1 (Trusted principle – Data Use stage of data lifecycle):*

The consolidated national dataset is a single source of truth. IHACPA teams access this dataset, apply their own processes and data transformations for different use cases. This is to obtain dataset(s) that are fit for purpose for each key use case. However, the documentation reviewed in relation to these team-specific uses and processes contained varying levels of detail and coverage on what each of the datasets are, what changes have been applied to the consolidated national dataset, and for what purpose. It is important to clearly and transparently document all adjustments made so that data limitations are well understood, and the appropriate dataset can be identified and applied to respond with confidence to ad hoc queries.

**Principle 2: Attainable**

> **Framework definition:** Data collection requirements are standardised, predictable and can be largely achieved using current and available data sources and structures

There are a number of enablers, systems, and tools used by IHACPA to ensure the data collection requirements are standardised and attainable by data providers. For example, the Data Request Specifications (DRS) defines the content, format, and permissible values for data and provides a standardised understanding between data providers on IHACPA's data expectations. IHACPA also consults with data providers to work towards agreement on a sufficient data submission.

No significant attainability gaps were identified as part of this review.

**Principle 3: Efficient**

> **Framework definition:** Data quality objectives and measures are determined in consideration of the resources required to achieve these outcomes and the materiality of the benefits.

There are examples of processes that IHACPA undertakes which consider the internal and external resources required to satisfy a requirement or accomplish a task. For example, validation reports generated by the data submission portal present results as a fatal error, critical error, or warning flag to communicate IHACPA's priorities with regard to data issues. This enables data providers to effectively allocate their resources to address areas of the greatest need within the timeliness constraints of the data submission.

*Principles Gap 2 (Efficient principle – Data Preparation stage of data lifecycle):* Stakeholder consultations identified some variation in the timing and completeness of jurisdictional responses on issues relating to their data submission.

**Principle 4: Innovative**

**Framework definition:** IHACPA implements new or improved approaches to data quality management to continuously improve outcomes.

In early 2025, IHACPA developed a NHCDC dashboard for jurisdictions' (NHCDC 2023-24) data submissions. The NHCDC dashboard is an innovative way to visualise and communicate data features of the cost and activity data internally and to data providers in a streamlined process. IHACPA is also reacting to disruptions in the health industry, including the post-COVID growth in virtual care. The policy and guidance documents reviewed focus mainly on IHACPA's responsibility to research innovative funding and costing models, and are therefore not displaying IHACPA's full involvement in innovation outside of these areas.

*Principles Gap 3 (Innovative principle – Data requirements stage of data lifecycle):*

Policy and guidance documents reference the importance of innovation, but do not communicate its purpose of vision, nor include areas where IHACPA anticipates disruptions and how it will manage change.

*Principles Gap 4 (Innovative principle – Data requirements stage of data lifecycle):*

IHACPA provides improvements in data quality management to a number of disruptions across the health landscape, however these activities were not found to be clearly captured under IHACPA's innovation goals. For example, the current Work Program and Corporate Plan states that IHACPA is undertaking a program of work to better understand virtual care in Australia, which can be considered an innovative opportunity to improve approaches to data quality management. This may result in limited stakeholder visibility of innovation being pursued.

**Principle 5: Transparent**

**Framework definition:** Clear consultation, expectations and communication between IHACPA, data providers and other stakeholders on what data is needed and available, in what form, for what purposes, and how it is validated, analysed and managed.

IHACPA provides data providers with documents such as the DRS, Three Year Data Plan, and Data Compliance Policy, outlining data requirements and expectations.

*Principles Gap 5 (Transparent principle – Data requirements stage of data lifecycle):*

Although some documentation was provided for the review which outlined team-specifications and data processes within IHACPA, no holistic documentation was provided which covered the end-to-end data lifecycle spanning across teams. Without regularly maintained and accessible documentation on the end-to-end data lifecycle, interdependencies and processes for the transfer and use of data between teams may not be fully understood across all key stakeholders. In light of

this, the process flow diagram was developed to present a holistic representation of IHACPA's data lifecycle and data flows between separate teams in IHACPA.

*Principles Gap 6 (Transparent principle – Data collection stage of data lifecycle):*

Data Quality Statements provided by data providers for the public and private hospital data collection include information on governance, summary of results, and compliance with Australian Hospital Patient Costing Standards (AHPCS) to varying degrees. As part of data governance, data providers outline the general quality assurance process that the data undergoes before submission. The Data Quality Statements vary in depth when describing the quality assurance process undertaken by data providers in their data provision, making it difficult for IHACPA to understand the potential limitations associated with data accuracy.

*Principles Gap 7 (Transparent principle – Data use stage of data lifecycle):*

IHACPA currently has a communications document to record interactions with external stakeholders regarding data quality and limitations. However, the review did not identify similar documents or databases that record internal interactions within IHACPA teams regarding data quality and limitations. Inclusion of such a document may enhance communication of data limitations when common datasets are analysed and used by different teams within IHACPA.

*Principles Gap 8 (Transparent principle – Data use stage of data lifecycle):*

The application of business rules, such as trimming procedures, are communicated to data providers through the release of technical documents. Whilst these documents outline how the checks are performed, not always outline the rationale of business rules and checks performed, or how these business rules should be updated over time for their continued use in subsequent data collections.

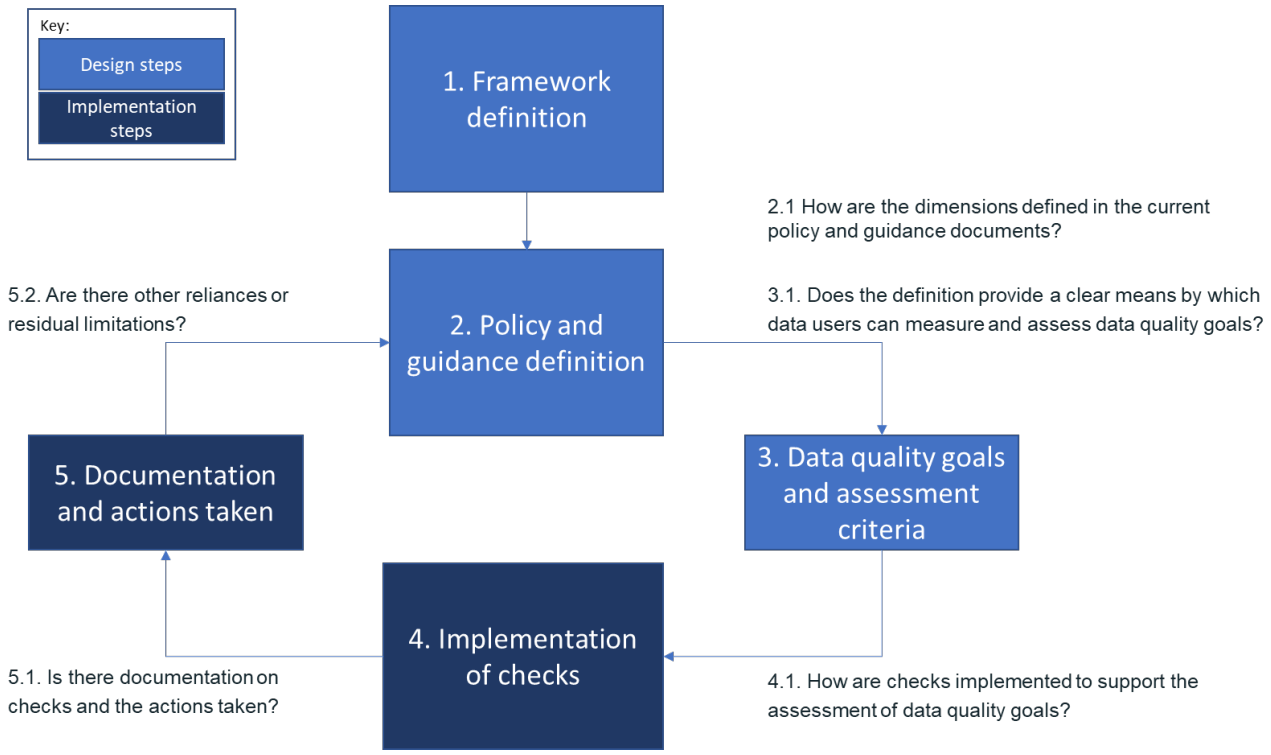
## 4.3 Data quality dimensions

IHACPA defines, measures and manages data quality through several dimensions: accuracy, coherence, completeness, confirming, coverage, suitability, timeliness and uniqueness. These dimensions (framework dimensions) are used to evaluate how informative and useful the data is for a given purpose.

This section covers the design and implementation steps of the framework dimensions across the data lifecycle. This includes observations on how each dimension is defined in policy and guidance documents and how IHACPA currently performs checking processes in line with the definition of each dimension, as well as an identification of any gaps. For more information on which policy and guidance documents each dimension is mapped to, see Appendix D: Mapping of policy and guidance documents to data quality dimensions.

The review of the framework dimensions focused on the high-level control cycle processes for designing and implementing definitions. The steps which comprise this process, and the associated considerations for each step, are outlined in Figure 3.

## Figure 3: High-level control cycle process for designing and implementing framework dimensions




Note: The numbering above is based on the sequential order of high-level control cycle process step, and number of considerations. For example, 5.2 relates to step 5: "Documentation and actions taken", and is the second key consideration for that step "Are there any other reliances or residual limitations?".

The control cycle process was used to identify gaps in how IHACPA is defining and implementing checks to assess data quality. Table 4 expands on steps 1 - 5 of Figure 3.

## Table 4: Control cycle steps for framework dimensions

| Design steps | | |
|---|---|---|
| **#** | **Step name** | **Description** |
| 1 | Framework definition | The definition is in the data quality definition section of the framework, and is used to measure the degree to which data is fit-for-purpose. |
| 2 | Policy and guidance definition | This is the definition that is communicated to stakeholders and the wider public through policy and guidance documents. This definition should align with what is in the framework. The control cycle process for this step should consider:<br>*2.1 How are the dimensions defined in the current policy and guidance documents?* |
| 3 | Data quality goals and assessment criteria | The policy and guidance definition should include measurable goals towards data quality, and a threshold to assess whether it |

| | | |
|---|---|---|
| | | is obtained. The control cycle process for this step should consider:<br>*3.1 Does the definition provide a clear means by which data users can measure and assess data quality goals?* |
| **Implementation steps** | | |
| **#** | **Step name** | **Description** |
| 4 | Implementation of checks | The measurable goals should be implemented as checks before the data is used. The control cycle process for this step should consider:<br>*4.1 How are checks implemented to support the assessment of data quality goals?* |
| 5 | Documentation and actions taken | The results of these checks should be documented and compared to threshold values to assess data quality. Documentation should also include any actions taken as a result of the checks, and residual limitations where checks were not applied. The control cycle process for this step should consider:<br>*5.1. Is there documentation on checks and the actions taken?*<br>*5.2. Are there other reliances or residual limitations in the data?* |

Table 5 summarises the framework dimensions gaps identified as part of the review.

### Table 5: Summary of framework dimensions gaps identified

| **Design steps: Policy and guidance definition, data quality assessment criteria** | | |
|---|---|---|
| **Gap** | **Gap description** | **Findings** |
| Dimensions Gap 1 | Multiple dimensions – policy and guidance definition step of control cycle | Based on the review, the extent to which the data quality dimensions are defined in IHACPA's current policy and guidance documents is variable. Several dimensions are explicitly defined, some are implicitly defined, and for some no clear definition was identified.<br>• Explicit definitions for timeliness, conformity and uniqueness were identified.<br>• Implicit definitions for coherence, coverage, and suitability were identified.<br>• No clear definitions for accuracy and completeness were identified. |

| | | |
|---|---|---|
| Dimensions Gap 2 | Multiple dimensions – data quality goals and assessment criteria step of control cycle | The description of data quality goals and explicit identification of assessment criteria is variable and, in some cases, missing.<br>• Both data quality goals and definitive thresholds or considerations against which data quality could be assessed were identified for *conformity*, *timeliness*, and *uniqueness*.<br>• Only data quality goals were identified for *coherence*, *coverage*, and *suitability*. However, for these dimensions no definitive thresholds or considerations against which data quality could be assessed were identified.<br>• Neither data quality goals nor definitive thresholds or considerations against which data quality could be assessed were identified for *accuracy* and *completeness*.<br>For example, the definition of *coherence* is implied by requesting variables that will be used for data linkage. This suggests that linkage between activity and cost data can be used as a goal to assess data quality. However, the definition does not include a threshold to assess data quality – such as a desired or expected linkage proportion between activity and cost data or what level of difference would require further consideration or action. |
| **Implementation steps: Implementation of checks, documentation and actions taken** | | |
| **Gap** | **Gap description** | **Findings** |
| Dimensions Gap 3 | Multiple dimensions – documentation and actions taken step of control cycle | IHACPA documents the outputs of checks it performs through validation reports and Quality Assurance (QA) reports. However, no documentation for *coherence* and *completeness* were identified outlining the conclusions reached or justification for subsequent action taken, or why no action was required.<br>Extending the example for *coherence*, IHACPA performs linkage checks between activity and cost submissions, and documents these checks in the validation reports in the data submission stage. However, there is limited detailed documentation around what constitutes an acceptable linkage result. Documentation associated with this check does not specify the thresholds that warrant escalation or further action to be taken (including when no action is required). |
| Dimensions Gap 4 | *Accuracy* dimension – documentation and actions taken step of control cycle: | For the public hospital data, IHACPA relies on the Statement of Assurance and Data Quality Statement to assess accuracy. However, the varied level of detail included in these documents, and IHACPA's limited visibility over the underlying processes of data providers, pose challenges for IHACPA to consistently conclude on data quality, in line with their requirements. |

The subsections below outline the specific findings within the design and implementation steps for each of the framework dimensions, as an extension of the gaps in the table above.

**Dimension 1: Accuracy**

**Framework definition:** The degree to which the data matches reality.

*Design findings:*

No clear definition of accuracy was identified in the policy and guidance documents reviewed. Although, for public hospital data, the policy and guidance documents contain an expectation for Statement of Assurance to include commentary on the steps taken to promote accuracy, they do not define a measure or threshold for accuracy. The absence of a clear definition and measures to assess accuracy may lead to inconsistencies in the interpretation of data accuracy between jurisdictions and IHACPA.

*Implementation findings:*

IHACPA has a limited ability to perform checks on accuracy without further data collection or a review of underlying records using methods such as an Independent Financial Review, as previously undertaken for public hospitals. As a result, IHACPA relies on the Statement of Assurance to define and assure data accuracy.

**Dimension 2: Coherence**

**Framework definition:** The data is consistent over time, stable across repeated processes or updates, and can be combined and compared with other data sources.

*Design findings:*

In the Addendum to the NHRA, coherence is referred to as ‘the consistent application and interpretation of data across systems.’ Coherence is also implicitly defined through the request of linkage variables in the DRS. This implies that completing the DRS is the goal to achieve data quality. However, limited guidance was identified relating to the threshold or considerations needed to assess data quality.

*Implementation findings:*

IHACPA performs linkage checks between activity and cost data to test coherence between the two data sources, as well as comparison checks using data submissions from prior years. Without a definitive threshold to assessing coherence in the data, it is difficult to consistently conclude on data quality, and difficult to comment on subsequent data actions.

**Dimension 3: Completeness**

**Framework definition:** The degree to which records are present.

*Design findings:*

The Data Compliance Policy does require from the Statement of Assurance a commentary on actions taken to promote data completeness. However, no explicit definition of completeness has

been identified in the policy and guidance documents, including a clear measure or threshold to assess whether data is complete.

*Implementation findings:*

The data submission portal checks for blank values in the data provided and documents this in the validation reports. QA reports are also used to perform and document checks. However, without a definitive threshold on completeness, it is difficult for IHACPA to consistently determine the level of data quality and comment on the actions taken to address data quality issues in the documentation.

**Dimension 4: Conformity**

**Framework definition:** The data adheres to the agreed syntax (format, type, range) of their definition.

*Design findings:*

The DRS explicitly define the syntax in which the data is expected. The Data Compliance Policy also explicitly states that costing data is to conform to the AHPCS.

*Implementation findings:*

The data submission portal checks the conformity of data submissions to the DRS, which is available in the validation reports.

**Dimension 5: Coverage**

**Framework definition:** The degree to which the data adequately covers the population or event (representativeness).

*Design findings:*

The Addendum to the NHRA defines coverage as “whole of system reporting” which implies the full population of data is required. The addendum also defines specific data coverage goals to increase coverage on “the representation of Aboriginal and Torres Strait Islander peoples’ experience”.

*Implementation findings:*

The NHCDC dashboard, for public hospitals, provides IHACPA with a comparison between jurisdiction-level data and the population data, as an aggregate of all jurisdictions. This allows IHACPA some ability to assess whether a submission is representative of the population. The checks are documented in QA reports which includes a comparison by care type and stream (e.g. Australian Refined Diagnosis Related Group (AR-DRG), Australian National Subacute and Non-Acute Patient Classification (AN-SNAP), etc.).

**Dimension 6: Suitability**

**Framework definition:** The extent to which the characteristics of the data meet the needs of the end use.

*Design findings:*

The DRS define the data IHACPA requires from data provider. As IHACPA has input into the ongoing development process for DRS, it is implied that IHACPA is making the necessary adjustments to their request to ensure that the data meet the needs of its end use. Although measures exist to assess conformity of data to the DRS, it is unclear what thresholds are used to consistently determine the overall suitability of the data submitted.

*Implementation findings:*

Beyond conformity to the DRS, IHACPA relies on the Statement of Assurance to state that the activity data is “official, complete, accurate, and fit-for-purpose at the time of submission”. The results of the DRS are recorded in the validation reports.

**Dimension 7: Timeliness**

**Framework definition:** The length of time between the availability of the data source and the event it describes.

*Design findings:*

The Three Year Data Plan explicitly defines the data submission dates and timelines, instructing that IHACPA is “provided data in the timeframes requested”. The penalty for non-compliance in public hospital data submissions involves being named in a published data compliance report and a notice from the Chief Executive Officer of IHACPA on the nature of non-compliance.

*Implementation findings:*

Timeliness is checked after the jurisdiction has submitted their data. A compliance report is created at the end of each quarter to document a jurisdiction’s compliance, for public hospital data.

**Dimension 8: Uniqueness**

**Framework definition:** There are no duplicate records, with only one record for each entity represented and each record is stored only once.

*Design findings:*

Unique data is not explicitly stated as a requirement in the NHRA but is strongly implied through the request for unique identifiers in the DRS.

*Implementation findings:*

The data submission portal checks for duplicates to ensure that the data submitted is unique. Residual limitations may exist in how the unique identifiers are created but this is addressed as part of accuracy in *Dimensions Gap 1 – policy and guidance definition.*

# 5. Recommendations

The project findings identified a number of opportunities for IHACPA to align processes across the whole data lifecycle with the foundational principles introduced in the framework. The findings also identified opportunities for IHACPA to improve its design and implementation of the framework dimensions across the data lifecycle. The recommendations outlined in Table 6 are intended to address all the findings outlined in Section 4.

Note: The gaps in the table of recommendations below correspond to the gaps outlined in Sections 4.2 and 4.3.

### Table 6: Table of recommendations

| **Recommendation 1: IHACPA to consider options to gain greater confidence on the accuracy of data provided by jurisdictions for the public hospital data submissions, either through independent review or by requesting greater levels of information from jurisdictions supporting their conclusions.** | |
|---|---|
| **Opportunities** | **Gap(s) addressed** |
| IHACPA should consider possible options to gain greater confidence on data accuracy. These options may include:<br>• conduct independent reviews or reconciliations of a representative sample of jurisdictional data to assess data accuracy. For example, via mechanisms such as the past used Independent Financial Review (IFR) or similar reconciliations, OR<br>• design a supplementary information request and feedback process that provides IHACPA with greater information related to adherence to the costing standards (AHPCS), jurisdictional data workflows, processes and methodologies, or other items that would otherwise be investigated as part of an IFR or other equivalent independent oversight mechanism for data integrity (as required under clause B76 of the NHRA). | **Principles Gap 6**<br><br>**Dimensions Gap 4**<br>*Accuracy* |
| **Recommendation 2: IHACPA to increase the completeness, coverage and holistic nature of documentation on internal data quality processes.** | |
| **Opportunities** | **Gap(s) addressed** |
| IHACPA should increase the completeness of documentation of internal data quality processes by taking the following steps:<br>• Explicitly document all elements of data quality checks and transformations performed within IHACPA. This should include the checks/transformations performed, their rationale/justification/purpose, the measures/ thresholds/considerations used to conclude on the results (if applicable), and the actions taken based on the results. Where documentation is currently implicit in processes (e.g. embedded in | **Principles Gaps 1, 5, 7 & 8**<br><br>**Dimensions Gap 3**<br>*Coherence, Completeness* |

| | |
|---|---|
| comments within analysis code) or currently does not exist, consider whether there is a need to explicitly incorporate this information in centralised documentation to improve accessibility.<br>• Develop and maintain process diagrams both within teams and across teams to cover the end-to-end data lifecycle and capture key handover points between teams. This includes maintaining and updating the process diagrams developed as part of this review.<br>Extend the scope of existing data issues and communications documents to be both external and internal to capture interactions within IHACPA teams regarding data quality and limitations. | |
| **Recommendation 3: IHACPA to review policy and guidance documents and revise (where required) to include explicit definitions, including measurable and actionable goals, for each data quality dimension.** | |
| **Opportunities** | **Gap(s) addressed** |
| IHACPA should review current policy and guidance documents and take the following steps:<br>• Include explicit definitions in policy and guidance documents for each data quality dimension. These definitions should include the criteria by which data quality in the dimension will be measured and assessed. The criteria, which may consist of measures, thresholds or goals, should be actionable to enable data providers and IHACPA to implement appropriate checks and report on results.<br>• Communicate these definitions to data providers ensure the definitions and their implications are well understood. | **Dimensions Gap 1**<br>*Accuracy, Coherence, Completeness, Coverage, Suitability*<br>**Dimensions Gap 2**<br>*Accuracy, Coherence, Completeness, Coverage, Suitability* |
| **Recommendation 4: IHACPA to investigate ways to improve the quality, timeliness and consistency of feedback and responses provided by data providers to queries on data quality. Options may include promoting the importance of feedback loops, formalising processes through existing committees, or other collaboration with data providers to work towards better practices on data quality processes and data integrity, as required under clause A154 of the National Health Reform Agreement (NHRA).** | |
| **Opportunities** | **Gap(s) addressed** |
| IHACPA should consider what options are available to make the feedback loop with jurisdictions more reliable and efficient. Possible options may include:<br>• using relationships between IHACPA and data providers to promote the importance of the feedback loops, and<br>• establishing formalised processes and requirements via existing committee structures (for example the TAC, JAC and NAC). | **Principles Gap 2** |

| Recommendation 5: IHACPA to clearly articulate its vision for innovation, integrating ongoing improvements in data quality with broader strategic activities in both pricing and non-pricing goals. | |
|---|---|
| **Opportunities** | **Gap(s) addressed** |
| IHACPA should clearly articulate its vision for innovation to ensure the goals reflect all of their activities by taking the following steps:<br>• Establish a vision, purpose and goals (including both pricing and non-pricing) for innovation in improving and advancing data quality.<br>• Review current and planned activities to ensure they are consistent with this vision, purpose and goals. For example, IHACPA's current program of work to better understand virtual care in Australia is an area of non-pricing innovation in data quality which could be assessed against IHACPA's vision, purpose and goals for innovation (once established). | **Principles Gaps 3 & 4** |

| Recommendation 6: IHACPA to review the framework on a regular basis. | |
|---|---|
| **Opportunities** | **Gap(s) addressed** |
| IHACPA should establish a protocol for review and update of the framework. This protocol should include both a regular timeframe for review as well as clearly defined trigger events that will initiate a review of the document outside of the regular review schedule. | Governance of framework |

# 6. Framework review methodology

## 6.1 Framework review methodology purpose

Disruptions to technology, legislative or policy changes, and changes to IHACPA's strategic focus or operating environment may have implications for IHACPA's data quality governance over time.

As a result, there is a need to refresh IHACPA's Data Quality Framework on both a regular basis and when key triggers of change are identified.

The framework applies to all datasets within IHACPA's portfolio of work. The framework appendices detailing the application of the framework to IHACPA's operations is currently focused on public hospital data, given the greater maturity of established processes. Private hospital data is well developed, however will need to be expanded, as well as other sectors that IHACPA operates in, such as aged care.

## 6.2 Framework sections

The framework documentation has been split into three sections, being the foreword, the framework itself, and the appendices dealing with the application of the framework. Each section requires a difference cadence of review over time.

**Foreword**

The foreword describes IHACPA's current strategic focus and operating environment. It is recommended to be updated annually to ensure its relevance.

**Framework**

The main body of the framework is designed to be an enduring asset, requiring less frequent updates than both the foreword and appendices. A review can occur under two scenarios:

Scenario 1:

The occurrence of specific events that initiate a targeted review of the framework to understand whether changes are required. These events may include (but are not limited to):

- changes in guidance issued by the Commonwealth Government
- significant changes to policy and legislation, affecting IHACPA's responsibilities, or
- significant changes to data governance and security practices.

The targeted review process involves firstly understanding the nature of the change that triggered the review, and then analysing whether it impacts the principles or definitions of the framework. If so, reaching consensus using stakeholder engagement and provider consultation on the implications would be required.

Otherwise, it may be more appropriate to reflect the change in the systems, enablers, or tools of the framework (and hence impacts the appendices).

Scenario 2:

A systematic review of the framework every 3 to 5 years is recommended to ensure the framework retains its relevance and benefits from any advances in data governance practice. The following is a high-level description of the activities to be undertaken:

- **Desktop review:** Conduct a comprehensive review of existing documentation, policies and procedures related to data quality. This should aim to identify any changes in regulatory requirements, standards, internal or external policies that may impact the framework.
- **Desktop research:** Identify improvements in practice related to data governance or data quality management both locally and internally.
- **Stakeholder engagement:** Perform consultations with committees, jurisdictions, and providers, to gather feedback on the current framework. This may include understanding changes in needs, new challenges that they are facing, and ways for improvement.

In addition to the above, perform a review of the considerations that were involved in the development of the current framework. Details are provided in the following appendices of this summary:

- Appendix A: Details of stakeholder engagement.
- Appendix B: Data quality objectives and requirements.
- Appendix C: National data collection and management principles and standards (including international approaches and best practice).

**Framework appendices**

The framework appendices document IHACPAs data quality stakeholders, data quality processes, and data quality enablers, tools and systems. A change to the appendices would occur under two scenarios:

Scenario 1:

An existing documented process in the appendices requires an update where:

- a data quality process improvement has occurred
- a trigger or disruption occurs requiring a change in systems, enablers, tools, or processes, or
- a change to the framework has been made.

The targeted review process involves firstly understanding the nature of the change above, and analysing where in the data process it impacts. If an impact is identified, consensus needs to be reached using stakeholder engagement and provider consultation on the required change.

Scenario 2:

The appendices are extended to include additional sector data. This would involve:

- **Desktop review:** Conduct a comprehensive review of existing documentation, policies and procedures related to data quality. This should aim to identify any changes in regulatory requirements, standards, internal or external policies that may impact the framework.
- **Stakeholder engagement:** Engage with key stakeholders, including branch representatives, jurisdictions, committees and working groups, to gather feedback on the current framework. Stakeholder consultations can be conducted through interviews, surveys, and workshops to understand their needs, challenges and suggestions for improvement. Their feedback should be incorporated into the framework to ensure it remains relevant and effective.
- **Process mapping:** Map out current data quality processes to identify any gaps, inefficiencies, or areas for improvement. If current processes have changed, ensure that the process maps are also updated to reflect any changes.
- **Software coding review:** Examine software code used in data processing and analysis to ensure it adheres to best practices and current standards. The code analysed should be validated for both accuracy and efficiency, making necessary updates to improve performance and maintainability. All changes made to the code should be documented for future reference.
- **Integrated status and feedback mechanism:** Implement a mechanism for continuous feedback among IHACPA and stakeholders. This feedback should be collected regularly and analysed to identify areas for improvement in data quality practices.

# Appendix A: Details of stakeholder engagement

## Stakeholder engagement approach

The stakeholder engagement approach for the development of the framework and the targeted review of quality assurance processes involved engagement with IHACPA staff and existing committees.

Additional stakeholder engagement activities will be conducted as part of the review and approval process for the framework.

## Engagement with IHACPA staff

**Branch representatives**

IHACPA staff expertise was represented by 12 nominated representatives from three branches: Pricing and Analytics, Hospital Policy and Classification, Costing and Data Infrastructure.

**Data Quality Challenges Workshop**

The engagement process commenced with workshops with the branch representatives focusing on current state data quality activities and challenges.

The workshops sought input on current state data collection, analysis and management processes, as well as the current state of data quality and data quality challenges for their teams.

Key findings from the workshops included:

- the importance of increasing the automation of data capture and validation
- impacts from hospital data evolutions, including new areas such as virtual care affecting consistency
- increased reporting from data providers that they are unable to provide the depth of data requested, reducing full datasets to representative samples
- the mismatch between actual data used and what data providers and external stakeholders expect data should represent
- setting transparent expectations regarding data quality
- checking consistency between data sources, and
- ongoing communication and feedback with stakeholders such as jurisdictions, private hospital groups, aged care providers and other government agencies.

### Individual meetings with branch representatives

Based on the findings of the workshops, follow-up meetings were scheduled providing an opportunity for a more in-depth discussion on particular aspects of IHACPA's approach to data quality.

The main topics of discussion for follow-up meetings were:

- current data management practices, including ingestion, cleansing and transformation
- quality assurance measures
- tool usage, including any current initiatives to review or update
- collaboration with relevant teams
- verification methods for jurisdiction data against quality criteria
- specific input and outputs from different processes
- communication with jurisdictions and other government agencies,
- future plans for enhancing data quality and processes, and
- verification of the process flow diagrams developed as part of the review.

## Engagement with committees

The National Hospital Cost Data Collection Advisory Committee (NAC) and Technical Advisory Committee (TAC) were identified in the stakeholder engagement plan to be consulted on data quality current state and challenges as part of developing the draft framework.

The main topics for discussion during IHACPA's consultation these committees are listed below, as well as a high-level summary of some of the key points raised by NAC and TAC representatives: data quality definition, including criteria that must be met prior to data submission.

All jurisdictions provided self-identified data quality definitions, however there was variation of definitions between jurisdictions. Most definitions mentioned accuracy, completeness, relevance, consistency, and timeliness. Some definitions also included coherence, accessibility, and interpretability.

### Current data collection practices

Data collection practices vary across jurisdictions. They typically comprise of a combination of the department of health directly extracting data from local hospital/health services' source systems and/or the submission of required data files by hospitals/health services to the department of health in line with pre-defined specifications.

### Processes for checking and treating outliers

Some jurisdictions indicated that identification and checking of outliers typically occurs via dashboards, QA reports, consultations and/or during the draft submission periods. The process includes outliers being reviewed, with an amendment and/or rationale provided by the source data provider (hospital or local health service) as needed.

### Quality assurance procedures for identification and handling of errors, including feedback mechanisms

Jurisdictions outlined a wide range of data quality procedures used for the identification and handling of errors. It was noted that checking and treatment processes occur at each level of the hospital, health service, and overall jurisdiction level. Where errors are found by the jurisdiction, there is a process of informing the health service for correction in their source systems, which may involve exclusion or modification if the source data cannot be corrected. Where data is unable to be reconciled, jurisdictions indicated they have procedures for excluding these records in the data submission.

**Processes in place to improve data quality**

Many jurisdictions advised they have continuous improvement processes in place to further improve data quality. Key activities noted include: workshops, working groups, training programs, targeted review areas, reflection exercises, audits, and comprehensive in-build checks.

**Annual review of methodologies including data collection, cost allocation and cost systems**

Reviews of costing systems setups and cost allocation methodologies are conducted annually by jurisdictions, as well as reviews of policies, processes and other documentation. Some jurisdictions also noted they conduct regular audits and reviews which involve the participation of key stakeholders such as patient costing managers. These reviews aim to check for accuracy and consistency and reduce the margin of error in data collection, cost allocation, and cost systems.

**Existing data quality assurance frameworks, processes or guidelines**

Jurisdictions indicated they have their own existing data quality assurance frameworks, guidelines and policies. It was noted that noted that many of these are based on national quality assurance checks, suggested practices from the Commonwealth, or internationally developed frameworks.

**Common drivers of data quality issues**

Jurisdictions identified a range of drivers affecting data quality. Key challenges include limitations of source data systems (such as the inability to link costed mental health data with the submitted activity data), staff turnover, the additional administrative burden that data collection imposes, and that the collection of data is of lower priority to the provision of quality clinical care.

**Feasibility of implementing a nationally consistent approach to quality assurance checks prior to submission to IHACPA**

Jurisdictions are supportive of the idea of a nationally consistent approach to data quality assurance checks. However, it was noted that such an approach would need to respect local data quality assurance processes and address the challenges of cost and potential rigidity that may affect each jurisdiction differently.

**Feedback on IHACPA's current processes, including the existing IHACPA Data Quality Assurance Framework and Data Request Specifications**

Jurisdictions highlighted the value of the DRS but suggested that there are improvements that can be made to acknowledge differences in data systems and reporting standards of hospitals/health services which may include consistently of terminology, application of validation checks, and the transparency of checks performed.

# Appendix B: Data quality objectives and requirements

## Current state data quality objectives and requirements

IHACPA assists the Australian Government to fund hospital and aged care services more efficiently by providing evidence-based price determinations and pricing advice. To fulfil this goal, IHACPA requires access to activity and cost data that is of a sufficient quantity and quality. Data quality is multi-dimensional, with quality relative to the needs of the users – a particular dataset may meet the quality requirements of one user, such as hospitals, but not another. Generic metrics for quality measurements cannot be applied directly to all datasets.

IHACPA relies on the provision of data from jurisdictions (state, territories and the Commonwealth Government) that is primarily collected for the delivery of healthcare and to make decisions about individual patient care. IHACPA's analysis is a secondary use of this data, which has varying levels of suitability for pricing purposes.

In the current model for determining ABF, each patient episode needs to be counted and costs need to be effectively allocated, including inpatient admissions, emergency department presentations, and outpatient appointments, as well as a range of mental health and rehabilitation services.

Assessing, determining, and analysing the impact of data quality is a key activity for IHACPA in fulfilling its main functions as an agency. IHACPA currently approaches managing data quality strategically through a number of mechanisms.

IHACPA documents and communicates its overall data quality assurance processes and requirements through the data quality assurance framework. However, the document has not been updated since its development and release in 2012, and is limited to hospital costing and activity data, as per the agency's scope of work at the time.

IHACPA has established formal mechanisms for consultation with jurisdictions and other stakeholders on data quality through the committee framework, as well as formal and informal discussions with other government agencies, such as the National Health Funding Body (NHFB) and Australian Institute of Health and Welfare (AIHW).

IHACPA sets standards and requirements for the data collected from stakeholders in the Data Request Specifications (DRS), which are published on their website. These standards and requirements are developed based on IHACPA's strategic objectives, what the agency needs to deliver for government and what jurisdictions can provide. These standards are also developed in consultation with jurisdictions and other health reform agencies. These specifications aim to abide by the principle of data rationalisation for all health reform agencies – "single provision, multiple use".

The data is collected through IHACPA's central portal. As jurisdictions submit the files, they are subject to a number of security checks and quality validation processes. Once validated, the data is stored in IHACPA's SDMS, which meets Australian Government security and privacy requirements.

For ABF data, jurisdictions must submit a statement of assurance in March and September, stating that the data is complete, accurate and fit for purpose at the time of submission. The statement of assurance is not made publicly available. For NHCDC data, jurisdictions are required to provide a Data Quality Statement (DQS) with their final data submission for the year and are published on their available. The DQS should outline information on governance, summary of results, and compliance with Australian Hospital Patient Costing Standards (AHPCS).

The data undergoes further quality assurance processes by IHACPA. Initially the data is profiled and validated against the DRS, using Redshift as part of the Extract, Transform and Load (ETL) processes. This is managed by the Data Acquisition team. The data is then used by different teams in IHACPA to complete their part of pricing determinations and advice for health funding. IHACPA teams perform additional quality assurance processes as part of their analysis, relevant to the specific calculations and outcomes, such as comparison with previous years' submissions.

IHACPA publishes Data Compliance Reports including a Quarterly Report, as required by the National Health Reform Agreement.  Dependent on their findings, IHACPA may also go back to jurisdictions with queries that require clarification and/or resubmission. There may also be discussion with other health reform agencies on specific data quality issues.

Future data requirements are communicated in advance to jurisdictions through the IHACPA Three Year Data Plan, which is released annually.

Finally, IHACPA has internal policies and frameworks that identify roles and responsibilities for managing data, including data quality, such as the Data Governance Policy.

# Future state data quality objectives and requirements

IHACPA's current key focus as an agency is to support public hospitals and aged care services to improve efficiency in, and access to, services through the provision of independent pricing determinations and advice and designing pricing systems that promote sustainable and high-quality care. Processes for collecting and analysing aged care data are a more recent function for IHACPA and are less mature than those for public hospitals. It is expected that these processes will be refined over time in collaboration with key stakeholders.

IHACPA's portfolio of work may also expand in the future. IHACPA also investigates other areas as requested, such as prostheses.

The provision of timely, accurate and reliable data is vital to IHACPA in fulfilling its legislated functions. IHACPA requires data that is of suitable quality to fulfill its mission, while accounting for the resources required from stakeholders to produce this data as part of delivering services.

IHACPA's desired future state incorporate the materiality of data collection with its purpose and plan as an agency. Data quality has a direct impact on decision making and allocating resources for health care for Australians.

IHACPA has six strategic objectives - as defined in the IHACPA Work Program and Corporate Plan 2024-25 - which can be linked to data quality target state and benefits, as shown in Table 7.

### Table 7: Strategic objectives and data quality target state

| # | Strategic objective | Data quality target state | Benefits |
|---|---|---|---|
| 1 | Perform pricing functions | **Accessible**: Data is available for analysis and modelling.<br>**Transparent**: Quality is clear and measurable.<br>**Reliable**: Data reflects reality to a known degree of confidence. | Increased systematic methodological rigour to data quality validation processes as part of the pricing functions |
| 2 | Refine and develop hospital and aged care activity classification systems | **Standardised**: Data adheres to known standards.<br>**Consistent**: Data is recorded and reported as per expectations. | Enables a nationally consistent method of classifying all types of patients, their treatment, and associated costs in order to provide better management, measurement and funding of healthcare services |
| 3 | Refine and improve hospital and aged care costing | **Measurable**: The limitations of data are known and quantifiable.<br>**Enhanced**: Data is able to be combined and integrated with other data sources. | Supports the agency to provide the strongest possible evidence for pricing determinations and advice in the most efficient way |
| 4 | Determine data requirements and collect data | **Compliant**: Data adheres to regulations and requirements, verified through quality checks.<br>**Coherent**: Data is the same as an overall target standard.<br>**Responsive**: Data is supplied within the agreed timeframe. | Increased innovation and accuracy in mechanisms designed to source, request and communicate on data needs |
| 5 | Investigate and make recommendations concerning cost-shifting and cross-border disputes | **Efficient**: Data collection and analysis can be performed with minimum viable resources.<br>**Timely**: Data is valid and accurate for the given time period. | Improved ability to rapidly respond to such requests and provide the necessary advice with a high degree of confidence |
| 6 | Conduct independent and transparent decision-making and engage with stakeholders | **Predictable**: Requests for data are known and standardised.<br>**Comprehensive**: Data is complete and representative. | Increased confidence in pricing determinations and advice, with explainable and traceable processes |

The agency's key performance indicators – outlined in the IHACPA Work Program and Corporate Plan 2024-25 - also provide further guidance on the role of data quality in enabling IHACPA to perform its key functions, as shown in Table 8.

### Table 8: Key performance indicators and data quality current state

| # | Key performance Indicators | Data quality current state |
|---|---|---|
| 1 | Support public hospitals and aged care services to improve efficiency in, and access to, services through the provision of independent pricing determinations and advice and designing pricing systems that promote sustainable and high-quality care. | The annual Pricing Frameworks are evidence-based, developed using robust and accurate data that is well-understood.<br>The determinations for public hospital service and annual pricing advice for aged care are evidence-based, developed using timely and comprehensive data from a range of sources. |
| 2 | Fulfill the reporting and performance requirements under the PGPA Rule. | Data submissions are assessed through quality assurance processes and the findings discussed with stakeholders for the purpose of continuous improvement and prioritising data quality activities<br>Accurate and reliable data is available for the purpose of IHACPA's corporate reporting |

IHACPA has identified three risk categories for the agencies, two of which can also impact the desired future state of data quality, as shown in Table 9.

### Table 9: Agency risk categories and data quality

| # | Risk category | Impact on data quality |
|---|---|---|
| 1 | Reputational | The role of IHACPA as an independent agency that produces reliable and accurate pricing determinations and advice relies on data quality. An improved understanding of data quality and a plan to improve data quality over time would support IHACPA producing outputs with increased credibility and dependability. IHACPA also continues to be trusted by stakeholders to collect, analyse and store this information over time. |
| 2 | Data and information governance | Improved data quality reduces the risk of inaccurate or delays to the provision of pricing determinations and advice. Reducing the resources required to profile, cleanse, check and consult on data submissions means that determinations and advice can also be produced more efficiently. |

Overall, IHACPA requires continuous improvement in their approach to data quality in order to achieve the agency's key objectives. Data is central to the agency's vision and operations. IHACPA

would benefit from a better understanding of the accuracy of data, faster provision of data and more assurance from stakeholders that data submissions are accurate.

# Opportunities

Some of the opportunities for IHACPA to improve data quality assurance are listed in Table 10.

### Table 10: Opportunities to improve data quality assurance

| # | Opportunity | Impact on data quality |
|---|---|---|
| 1 | Improved understanding of the impact of data quality on outcomes | Improve understanding the impact of data quality issues on IHACPA's strategic outcomes, namely the robustness of pricing determinations and advice.<br><br>This includes more specific measures for data quality issues and impacts, and potential costs for the Australian Government and stakeholders for less accurate pricing determinations and advice. |
| 2 | Identify critical data requirements and those which support or guide. | Prioritise aspects of data submissions, identifying critical data requirements and those which support or guide.<br><br>This distinction can assist jurisdictions in focusing their data quality activities, and may result in changes that could be implemented relatively easily by jurisdictions that would have an outsize positive impact on data quality. |
| 3 | New approaches for jurisdictions to understand and communicate the quality of data submissions | In collaboration with jurisdictions, identify new approaches for jurisdictions to understand and communicate the quality of data submissions, for IHACPA to incorporate into their determinations process. |
| 4 | Support increased sharing and collaboration across jurisdictions on quality assurance | Some jurisdictions have expressed a desire for further information sharing and collaboration across jurisdictions, supported or even facilitated by IHACPA. While jurisdictions have different delivery scales and models, there is an opportunity for knowledge transfer and upskilling in a specialised area of expertise. Lessons learned from problems investigated, addressed and solved would be of particular utility. |
| 5 | Consider alternate sources of data to be used in modelling and analysis, and automated systems in data collection and validation | Investigate alternate sources of data to be used in modelling and analysis. There are ongoing improvements and increased adoption of digital health technologies, such as the increased take-up of ePrescribing. These technologies are a potential alternate source of data for IHACPA to explore.<br><br>Increasing the emphasis on automated systems in data collection and validation is also likely to reduce errors. Data quality will remain a moving target, as ongoing changes to |

| # | Opportunity | Impact on data quality |
|---|---|---|
| | | healthcare services and delivery impact the data collected by IHACPA. |
| 6 | Consider alternate methodologies for pricing determinations that require less high-quality data | Investigate alternate methodologies for pricing determinations that require less high-quality data.<br><br>For example, explore potential options and impacts from adopting a representative sampling approach to data submissions. As the complexity of the data collection increases, remaining flexible and nimble in how agency outcomes are achieved could result in higher quality data and in turn pricing determinations. |

# Appendix C: National data collection and management principles and standards

A review of national data development principles, other data collection practices and frameworks across government assisted in forming a view on leading practice in government data collection and management. Relevant and applicable findings for IHACPA data collection and management have been incorporated and referenced in the updated framework.

## Australian Government positions and requirements

This section documents the relevant Australian Government legislation, frameworks and policies relating to data governance, collection and management. The updated IHACPA framework incorporates these positions and requirements.

### Table 11: Relevant legislation relating to data governance, collection and management

| Relevant legislation relating to data governance, collection and management | |
|---|---|
| *National Health Reform Act 2011* | |
| Administered by | Department of Health, Disability and Ageing |
| Context | The Act governs the role, functions and responsibilities of the national health reform bodies, including IHACPA. |
| Requirement / Position | The Pricing Authority has the function of determining data requirements and data standards to apply in relation to data to be provided by states and territories. This includes data and coding standards to support uniform provision of data, and requirements and standards relating to patient demographic characteristics and other information relevant to classifying, costing and paying for public hospital services. |
| Relevance to IHACPA | IHACPA can set data requirements and standards for jurisdiction submissions. |
| *Privacy Act 1988* | |
| Administered by | Office of the Australian Information Commissioner |

| | |
|---|---|
| Context | The Privacy Act regulates the handling of personal information about individuals. This includes the collection, use, storage, and disclosure of personal information for both government agencies and the private sector.<br><br>The Australian Privacy Principles (APPs) with the Act provide a comprehensive framework for the ethical handling, storage and use of personal data, ensuring that individual's privacy is respected and protected. |
| Requirement / Position | The Privacy Act places significant emphasis on the governance of personal information, particularly sensitive health data. The Act mandates that personal information can only be collected with consent and for purposes directly related to agency's functions. IHACPA must adhere to the Act and the Australian Privacy Principles (APPs), which establish guidelines for managing personal information, including in healthcare. This involves ensuring transparency in how personal information is handled. |
| Relevance to IHACPA | IHACPA must maintain the confidentiality and integrity of collected data, implementing strict data security measures to safeguard personal information against unauthorised access and breaches. In addition, IHACPA must incorporate the APPs into information-related policies and frameworks to ensure that personal information is collected and managed in a way that respects individual privacy and complies with national standards. |
| *Public Governance, Performance and Accountability (PGPA) Act 2013* | |
| Administered by | Department of Finance |
| Context | The PGPA Act establishes a coherent system of governance and accountability for public resources. It describes how Australian government agencies should manage their resources. |
| Requirement / Position | Australian Government data assets are agency resources, and therefore need to be managed responsibly and appropriately, with consideration for their overall value. |
| Relevance to IHACPA | As a data-driven agency, the PGPA Act establishes rules for the broader governance, performance and accountability of how IHACPA handles its data holdings. |
| *Data Availability and Transparency Act 2022* | |
| Administered by | Office of the National Data Commissioner (ONDC) |
| Context | The Act establishes a scheme for authorised sharing of public sector data by Australian Government agencies with accredited users – the DATA Scheme. |

| | |
|---|---|
| Requirement / Position | Australian Government bodies, known as Data Custodians, can share data with accredited users. Data sharing must be in the public interest, and cannot be shared for the purpose of enforcement or compliance. Data Custodians have no duty to share, but must provide reasons if refusing a data sharing request. The Data Sharing Principles manage the risk of data sharing by applying controls, and must be applied at all times. |
| Relevance to IHACPA | IHACPA has indicated in the Three Year Data Plan that it plans to review opportunities to share data under the Act in the future. The Act may also assist in IHACPA in gaining access to data for new functions or requirements in the future. |
| *Archives Act 1983* | |
| Administered by | National Archives of Australia |
| Context | The Act makes Australian Government agencies responsible for the management of Commonwealth records and for following records management requirements. |
| Requirement / Position | The destruction, transfer or alteration of Commonwealth records is subject to the National Archives' authorisation. |
| Relevance to IHACPA | Considering the broad definition of a Commonwealth record in the Act, any activities that involve the destruction, transfer or alteration of data by IHACPA as part of managing data quality need to consider the requirements under appropriate records authorities or other official guidance. |

## Table 12: Relevant Australian Government policies and frameworks

| Relevant policies and frameworks relating to data governance, collection and management | |
|---|---|
| Addendum to National Health Reform Agreement (NHRA) 2020-26 | |
| Administered by | Department of Health, Disability and Ageing |
| Context | The Addendum to the NHRA is an agreement between the Australian Government and all state and territory governments. It aims to improve health outcomes for all Australians and ensure a sustainable health system. It aims to drive best practice and performance using data and research. |

| | |
|---|---|
| Requirement / Position | The Addendum to the NHRA states that Commonwealth funding for public hospital services and functions under the addendum is dependent on the provision of data requested by the national bodies. It outlines principles for national bodies to follow in determining their data requirements.<br><br>The addendum outlines the arrangements for sharing information between jurisdictions and national bodies, emphasising the importance of timeliness and transparency for reporting and determinations. It states that jurisdictions will provide the national bodies with the data determined necessary to carry out their functions in accordance with their data plans.<br><br>The addendum commits jurisdictions to working together and with national bodies to share and work towards best practice approaches to data quality and integrity. Jurisdictions are responsible for data integrity within their systems and provide independent oversight mechanisms. Improvements to data quality and information available to inform clinicians' practice should reduce preventable poor quality patient care.<br><br>Finally, the addendum identifies the processes and guardrails for data matching. |
| Relevance to IHACPA | The addendum provides IHACPA with the authority and agreement to set standards, collect data and work in collaboration with jurisdictions to manage data quality and integrity. The current addendum has a limited focus on data quality, which could be expanded upon in future iterations. |
| Data and Digital Government Strategy | |
| Administered by | Digital Transformation Agency |
| Context | As part of establishing data and digital foundations, data will be managed as a valuable national asset. Government agencies will establish and invest in appropriate mechanisms, infrastructure and practices to support data curation, storage, protection and use. |
| Requirement / Position | Relevant actions include adopting best practice data collection and use to create data assets that support policy development and decision making, and embedding data standards into all data asset management functions, focusing on data quality: accuracy, completeness, auditability, consistency and timeliness. |
| Relevance to IHACPA | The strategy's emphasis on high-quality data management directly supports IHACPA's goals of ensuring accurate, complete, and timely data for effective health and aged care policy development and decision-making. |
| Policy for the Responsible Use of AI in Government | |
| Administered by | Digital Transformation Agency |
| Context | The policy aims to ensure that government plays a leadership in embracing AI for the benefit of Australians while ensuring its safe, ethical and responsible use, in line with community expectations. |

| | |
|---|---|
| Requirement / Position | In-scope agencies must designate an accountable official for the agency's use of AI. In-scope agencies must also make publicly available a statement outlining their approach to AI adoption and use. |
| Relevance to IHACPA | IHACPA has designated an AI Accountable Officer and has also implemented and published a publicly available AI Transparency Statement which outlines IHACPA's the approach to AI use. |
| APS Data Capability Framework | |
| Administered by | Australian Public Service Commission |
| Context | The framework establishes the language used to define capability areas in the APS across the data lifecycle. |
| Requirement / Position | The capability 'Data Quality' is defined as applying measures to ensure that data being used or produced is fit for purpose. Three proficiency levels are defined – Foundation, Intermediate and Advanced. Associated capability areas include Improvement and innovation, Data collection methodology, Integrate data, Exploratory data analysis and Specialist data analysis. |
| Relevance to IHACPA | Assists IHACPA in documenting the capabilities required as part of the data management and governance function. Allows for IHACPA to communicate with other government agencies and internal staff on the capabilities required to implement the framework. |
| Framework for Governance of Indigenous Data | |
| Administered by | National Indigenous Australians Agency |
| Context | The framework provides a stepping stone towards greater awareness and acceptance by Australian Government agencies of the principles of Indigenous Data Sovereignty. |
| Requirement / Position | It aims to provide Aboriginal and Torres Strait Islander people greater agency over how their data are governed within the APS so government-held data better reflects their priorities and aspirations. The framework applies to Indigenous data held by APS agencies. The framework guidelines include building data-related capabilities and providing knowledge of data assets. |
| Relevance to IHACPA | IHACPA holds data about Aboriginal and Torres Strait Islander peoples. Some aspects of this framework may apply for IHACPA. |
| Protective Security Policy Framework | |
| Administered by | Department of Home Affairs |

| | |
|---|---|
| Context | The PSPF helps Australian Government entities to protect their people, information and assets. Agencies apply the PSPF using a security risk management approach, allowing them to apply it in a way that best suits their individual security goals and objectives, their specific risk and threat environment, as well as their risk tolerance and security capability. |
| Requirement / Position | The PSPF controls access to Australian Government information, handling and storing arrangements to guard against information compromise, and safeguarding ICT systems to support the secure and continuous delivery of government business. |
| Relevance to IHACPA | IHACPA must align its security practices with the PSPF to protect against threats to its data assets. In addition, IHACPA must adhere to government-wide standards to safeguard the integrity and confidentiality of the data it handles. |
| Information Security Manual | |
| Administered by | Australian Cyber Security Centre |
| Context | The ISM outlines a cyber security framework that an organisation can apply, using their risk management framework, to protect their technology systems, applications and data from cyber threats. |
| Requirement / Position | The ISM offers comprehensive strategies and best practices for agencies to protect their digital assets, emphasizing robust security, regular assessments and ongoing enhancement of cybersecurity measures, including governance, risk management and incident response protocols |
| Relevance to IHACPA | IHACPA must ensure that its cybersecurity measures are in line with the ISM to protect against data breaches and cyber-attacks. |
| Building Trust in the Public Record | |
| Administered by | National Archives of Australia |
| Context | The policy emphasises the need for Australian Government agencies to hold authentic, complete and reliable information to make evidence-based decisions, provide sound advice and develop good policy. |
| Requirement / Position | The policy requires that Australian Government agencies manage information assets strategically with appropriate governance and reporting. Recommended activities to achieve this outcome include having up-to-date governance arrangements for all information assets, identifying information assets and registering them, and staff having the necessary skills and knowledge to manage information according to its value. |
| Relevance to IHACPA | Ensuring IHACPA's data governance arrangements are in place to support the implementation of the framework. |
| Australia's Third Open Government Partnership National Action Plan | |

| | |
|---|---|
| Administered by | Attorney-General's Department |
| Context | The Open Government Partnership is a multilateral initiative that aims to secure commitments from governments to promote transparency, empower citizens, fight corruption and harness new technologies to strengthen governance. |
| Requirement / Position | The plan includes the Australian Government's commitment to create transparency in the use of automated decision making and AI. It is intended to be achieved through improved governance, capability and guidance to support the safe and responsible use of AI in Australia. |
| Relevance to IHACPA | If IHAPCA pursues increased automation of data processing and analysis, this will involve ensuring that there is transparency about its use. |
| Foundational Four | |
| Administered by | Office of the National Data Commissioner (ONDC) |
| Context | Provides guidance for agencies on how they can improve their data practices and address the technical and cultural challenges that can limit their ability to get the most out of their data. |
| Requirement / Position | The guidance sets out four foundational data practices for this purpose and can help more mature agencies to reflect on the next steps in improving data capabilities. The practices include having a senior leader accountable for data, developing a strategy, implementing data governance mechanisms, and identifying and recording data assets.<br><br>To support discoverability of data, the ONDC also has guidance on developing a data inventory, a standardised list of data assets held by the agency. A data inventory supports the use and re-use of data, as well as ensuring it is adequately protected. A data inventory also supports agencies to meet accountability for data under the Act, as well as other legislation such as the PGPA Act, Privacy Act and Archives Act. |
| Relevance to IHACPA | Ensuring foundational practices, such as a register of data assets or data inventory, are in place will support IHACPA to continuously improve data quality practices over time. |
| Notifiable Data Breaches Scheme | |
| Administered by | Office of the Australian Information Commissioner (OAIC) |
| Context | The notifiable data breaches scheme requires organisations to notify individuals and the OAIC about significant breaches of personal information that are likely to result in serious harm. |
| Requirement / Position | Any agency the Privacy Act covers must notify affected individuals and the OAIC when a data breach is likely to result in serious harm to an individual whose personal information is involved. A data breach occurs when personal information an agency holds is lost or subjected to unauthorised access or disclosure. |

| | |
|---|---|
| Relevance to IHACPA | IHACPA must ensure it has processes in place for detecting, assessing, and responding to data breaches in compliance with the notifiable data breaches scheme. |

# National approaches and best practice

While other jurisdictions may not have a similar pricing determination and advice model, there are lessons to be learned from other jurisdictions' approaches to collecting and managing data, particularly health data collection. Similarly, other government agencies in Australia have identified best practice approaches.

In this section, Australian and international examples of frameworks, policies and guidance on data quality in the government and/or health sector were reviewed, particularly focusing on use of health data for decision-making and policy development.

IHACPA can consider some of the key findings from this review, which have been incorporated into the development of the updated framework.

The documents reviewed are listed in Table 13 and Table 14.

### Table 13: Australian documents reviewed

| Document title | Jurisdiction | Date released / updated |
|---|---|---|
| The ABS Data Quality Framework | Commonwealth | 2009 |
| Australian Institute of Health and Welfare (AIHW) Data Governance Framework | Commonwealth | October 2022 |
| Data Quality Compliance in the Consumer Data Right - Discussion Paper | Commonwealth | October 2022 |
| NSW Government Data Quality Reporting Tool | NSW | Unknown |
| Victorian Government Data Quality Statement Template | Victoria | March 2018 |
| NSW Health Data Quality Assurance Framework for Activity Based Management | NSW | July 20216 |
| NSW Government Information Management Framework | NSW | 2018 |
| NSW Government Standard for Data Quality Reporting | NSW | October 2015 |

### Table 14: International documents reviewed

| Document title | Jurisdiction | Date released / updated |
|---|---|---|
| Canadian Institute for Health Information (CIHI) Information Quality Framework | Canada | 2017 |
| Data Quality Framework for EU Medicines Regulations | European Union | October 2023 |
| Data Toolkit | New Zealand | 2021 |
| Data.europa.eu Data Quality Guidelines | European Union | August 2021 |
| Education Quality and Accountability Office's (EQAO) Data Quality Framework | Canada | November 2020 |
| European Health Data Space Data Quality Framework | European Union | May 2022 |
| European Statistical Systems (ESS) Handbook for Quality and Metadata Reports | European Union | November 2021 |
| Eurostat Quality Assurance Framework | European Union | 2008 |
| The Government Data Quality Framework | United Kingdom | December 2020 |
| GS1 Data Quality Framework | Global | October 2010 |
| Guidance on Data Quality | Canada | January 2024 |
| Guidelines for the Template for a Generic National Quality Assurance Framework (NQAF) | Global – United Nations Statistical Commission | February 2012 |
| Health Data Research UK Data Utility Framework | United Kingdom | November 2020 |
| Ministry of Justice Data Quality Framework | New Zealand | June 2008 |
| National Health Service (NHS) Data Quality Assurance Framework for Providers (Part 1 and 2) | United Kingdom | January 2020 |
| Office for National Statistics Data Quality Management Policy | United Kingdom | Unknown |
| Quality Assurance Framework of the European Statistical System | European Union | 2019 |
| Towards European Health Data Space (TEHDAS) Recommendations on a Data Quality Framework for the European Health Data Space for Secondary Use | European Union | September 2023 |

### Purpose

Data-driven organisations emphasise that quality is a key driver in all activities and at all stages of the data lifecycle. For example, the Canadian Institute for Health Information (CIHI) states “quality is at the heart of everything [we do]. It is embedded in our vision and mandate.” The need for trusted information is paramount, critical to the agency’s ongoing success and relevance.

For health agencies like the UK’s National Health Service (NHS), developing a framework for data quality is about providing a focal point for the sharing of data quality assurance best practice across both the organisation and providers. This framework aims to ensure that clinical and administrative systems are configured to maximise data quality at the point of capture, with staff suitably trained to meet this goal.

### Scope

Policies and frameworks for data quality apply to different parts of an organisation’s work, and may be broader or narrower depending on the agency’s purpose. The UK Office for National Statistics Data Quality Policy, for example, applies to not just data acquired or collected, but to any intermediate data product necessary to the data lifecycle.

Other agencies consider maturity models for managing data quality, with higher maturity levels supporting the strongest possible evidence in the most efficient way.[2]

### Definition

Many organisations include the concepts of “fitness-for-purpose” and/or “reflecting reality” in their definition of data quality. Some frameworks cite the precise ‘fitness’ for the context, such as fit for purpose for users’ needs in relation to health research, policy making and regulation and the data reflect the reality which they aim to represent.[3] Fitness-for-purpose also implies an approach that includes both elements of technical quality and utility.

### Dimensions

Most organisations consider data quality to be multi-dimensional, assessing it against a range of criteria. For some, data quality dimensions need to be measurable – the metrics must be trackable, quantifiable and auditable.[4] Quality dimensions are not mutually exclusive and need to be balanced against one another to best meet users’ needs. Sometimes improvements in one dimension can lead to deterioration in another – trade-offs are often necessary, such as a reduction in accuracy in order to improve timeliness.[5] Many organisations prioritise data quality dimensions that align with their user and business needs.[6]

Data quality dimensions frequently used in policy documents include:

- Accuracy and/or reliability
- Completeness

[2] Data Quality Framework for EU Medicines Regulations
[3] European Health Data Space Data Quality Framework
[4] Canada’s Education Quality and Accountability Office’s (EQAO) Data Quality Framework
[5] CIHI’s Information Quality Framework
[6] UK Government Data Quality Framework

- Uniqueness
- Consistency
- Timeliness
- Validity
- Relevance
- Accessibility and clarity
- Coherence and comparability
- Interpretability
- Extensiveness
- Coverage

Not all frameworks use ‘accuracy’ as a data quality dimension, due to the difficulty in measuring it particularly when collected from multiple providers for secondary uses. For example, the European Health Data Space Data Quality Framework only includes the measurable dimensions of reliability, relevance, timeliness, coherence, coverage and completeness.

**Activities**

Generally, data quality initiatives should focus on continuous improvement, encouraging good practice, design, development and implementation of toolkits for quality assurance. Resources should be allocated to support data quality-focused work.[7]

The Eurostat Quality Assurance Framework describes three categories of tools for improving data quality to be used according to specific aims:

- Documentation and measurement – the complex information obtained from measurement and documentation is selected and structured. Methods and tools, like identifying key process variables and quality indicators are used in the individual domain.
- Evaluation – with quality work evolving over time, evaluation goes a step further and data is evaluated against internal or external standards. Quality assessments are conducted, with the improvement actions and identification of good practices as the outputs.
- Conformity with recognised standards – Labelling further condenses the information to demonstrate the compliance with defined standards and requirements, helping to enhance trust and credibility in the outcomes.

Any remediation activities should be documented to include the impacts to the rest of the data lifecycle. Changes or remediation of data quality should be communicated to all users within the lifecycle.[8]

Data quality controls can be implemented differently depending on if the “true facts” (the original source records) are known and accessible for validation. However, even if the source records are available, the validation process can be costly and time-consuming. Instead, data can be tested via intrinsic plausibility metrics, and specifically by assessing the data in respect to:

- other data in the same data set – the test would detect logical or factual contradictions

[7] European Health Data Space Data Quality Framework
[8] UK Office of National Statistics Data Quality Policy

- external reference ranges – some measured quantities cannot exceed a certain magnitude, or
- plausible trends – certain data can be validated when observed individually, but the collective trend of all data of a kind should follow expected distributions or trends.[9]

**Limitations**

When assessing data quality, there are other considerations that must be taken into account, such as cost, resource availability, collection burden, privacy, confidentiality and security. It may not be reasonable or even desirable to seek higher-quality data under all circumstances.

Some jurisdictions distinguish between primary use and secondary use of health data. Primary use of health data is when health data is used to deliver healthcare and to make decisions about the care of the individual from whom it was collected. Secondary use of health data is the use of aggregated health data from population-level sources to improve personal care planning, medicines development, safety monitoring and policy development.[10]

While secondary use of health data has important outcomes and benefits, Towards European Health Data Space (TEHDAS), a specialist group funded by the European Union, concludes that "it is not an option to compromise the quality of patient treatment to improve the quality of data for secondary use. Fitness for secondary use is a secondary incentive."[11] Health data quality recording should focus on its primary use in patient treatment, and be as clinically relevant and meaningful as possible. Standards, policies or guidelines aimed at improving secondary use of data should be introduced cautiously, and should not draw resources away from the point of care.

Finally, it is worth noting that data quality can be related to incentives – TEHDAS also notes that procedures or diagnoses recorded for reimbursement may have better coverage and completeness.

**International Standards**

Two international standards were identified across data quality frameworks:

- ISO 9000 Quality management system: Covers the basic terms and concepts of quality management systems.
- ISO 25012 Software engineering – Software product quality requirement and evaluation – Data quality model: Defines a general data quality model for data retained in a structure format within a computer system. Defines quality characteristics for target data used by humans and systems.

**Culture**

The importance, impact and involvement of organisational culture on data quality is emphasised in many policies and frameworks. Some of the key mechanisms of a quality culture include training, continuous improvement and communications.

[9] Data Quality Framework for EU Medicines Regulations
[10] The Open Data Institute (UK)
[11] Recommendations on a Data Quality Framework for the European Health Data Space for Secondary Use

For CIHI, quality is a feature of organisational culture – a successful and effective culture of quality must be communicated, nurtured and reinforced at every level, at every opportunity.

Training is frequently raised as key to implementing a quality culture:

- A training program instils and maintains a quality culture throughout the organisation.[12]
- A quality culture is spread in the organisation by means of regular training programs supporting the implementation of the quality policy, training on the job, regular training courses, workshops and other initiatives.[13]
- All employees have a role to play in ensuring quality, and are trained to fulfill these roles.[14]

**Outcomes**

The benefits of managing and improving data quality are varied, based on the organisational remit and objectives.

Some of the expected benefits from data quality assurance activities potentially relevant to IHACPA include the following:

- Ensure easier surveillance of patient safety, allowing clinicians to use data confidently to drive local quality improvement initiatives and support new technology for direct care including artificial intelligence.[15]
- Increased trust in the organisation through increasing both transparency and the awareness for the need for continuous consultation, growth and improvement. As well, transparency around any limitations and actions taken to address those limitations engenders a sense of integrity, honesty and openness.[16]
- A better understanding of the data quality problem through data quality work that is proactive, evidence-based and targeted. While there is no such thing as 'perfect quality' data, government must strive for a culture of continuous improvement. Through improved management of data, government can achieve the high-quality data needed to deliver better outcomes for society.[17]

# Potential options and future changes

**The next iteration of the National Health Reform Agreement**

The current iteration of the National Health Reform Agreement has been extended for an additional 12 months and now expires at the end of 2026. Negotiations for the next iteration are currently underway. The next iteration provides an opportunity to make changes to improve data quality for developing price determinations, pricing advice and other end uses described in the addendum.

[12] ESS Handbook for Quality and Metadata Reports
[13] ESS Quality Assurance Framework
[14] CIHI's Information Quality Framework
[15] NHS Digital Data Quality Assurance Framework for Providers
[16] CIHI's Information Quality Framework
[17] UK Government Data Quality Framework

Based on consultation with jurisdictions as part of drafting the framework, there is scope and broad support for increasing the specificity of data quality assurances provided to IHACPA by data providers.

Jurisdictions could provide the results of detailed checks and assurances as part of submissions, providing further assurance and more conformity of data.

**New AI regulations for Australian Government agencies**

There are increasing opportunities for the use of new technologies such as in collecting, assessing and managing data quality.

With an increase in uses and options for new technologies such as Artificial Intelligence (AI), the Australian Government is looking to assure the responsible application of these technologies.

The Digital Transformation Agency (DTA) has released the Policy for the Responsible Use of AI in Government, with two requirements:

- Designate an accountable official for the agency's use of AI, if not already done so, by 30 November 2024.
- Transparency statements current and future use of AI must be in place by 28 February 2025.

As a Corporate Commonwealth Entity, this policy does not apply to IHACPA – however, out-of-scope agencies are still encouraged to comply with the policy. Doing so would support broader stakeholder acceptance of the use of AI by IHACPA.

IHACPA has developed their own internal policy for the responsible use of artificial intelligence. The policy outlines IHACPA's ethical and effective use of AI technologies, ensuring they align with IHACPA's values, productivity, and the security of the SDMS and other environments.

**Updated privacy legislation**

The first tranche of reforms to the Privacy Act were introduced to Parliament in September 2024 under the *Privacy and Other Legislation Amendment Bill 2024.* The bill intends to modernise privacy regulations in Australia and includes a number of points that IHACPA may wish to consider such as:

- a new statutory tort for serious privacy breaches
- enhanced data security obligations, and
- greater transparency for individuals regarding automated decisions that affect them

While these new points may not directly affect IHACPA's obligations at this time, the bill reflects an overall trend of growing community and government expectations for the management of personal information. Scrutiny of data management practices will likely continue to increase in the future for all data holders.

A second tranche of more substantial changes is expected to be announced in 2025, followed by further consultation with stakeholders.

While IHACPA's Privacy Policy outlines its activities and obligations for personal information, requirements for practices, procedures and systems for personal information, including

depersonalisation of data, secure storage and responses to data breaches, may change in the near future.

**New pathways to appropriate data sharing**

To support innovation and collaboration across different sectors, IHACPA can further promote access to its data holdings for appropriate purposes such as research and policy. Increased use of the data sets provides increased return for all stakeholders of the resources required to supply, collect, profile, cleanse, manage, store and archive datasets within the data lifecycle.

Currently IHACPA recognises that access to high quality and nationally consistent health information is essential for the conduct of research and analysis. IHACPA can release protected Pricing Authority data as per the Data Access and Release Policy, which incorporates the requirements set out in the addendum and relevant legislation for data access.

The DATA Scheme, introduced as part of the *Data Availability and Transparency Act 2022*, provides new options for both sharing and receiving datasets for IHACPA. Increasing the valid use of collected data increases the potential benefits to the Australian community of IHACPA's data collection, assisting the case for jurisdictions to invest in providing higher quality data.

IHACPA may also benefit in the future with access to datasets through the Scheme, providing an opportunity to explore new data sources and collection methodologies.

# Appendix D: Mapping of policy and guidance documents to data quality dimensions

This section presents observations from a targeted review of IHACPA's policy and guidance documents and maps them to the data quality dimensions that were introduced in the framework.

Note: "-" indicates that no mapping between the policy and guidance documents and the data quality dimension was identified during the targeted review.

## Table 15: Findings and observations when mapping data quality dimensions to policy and guidance documents

| Policy and guidance document | Data quality dimensions | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Accuracy | Coherence | Completeness | Conformity | Coverage | Suitability | Timeliness | Uniqueness |
| National Health Reform Agreement and Addendum | Identifies a reliance on the Statement of Assurance. | States the importance of structured and consistent data collection methods. | Identifies a reliance on the Statement of Assurance. | Provides examples of conformity and fit for purpose in relation to other Australian Government practices. | Sets the aim to increase coverage of reporting into primary care.<br>Implied by "whole of system reporting". | Encourages the development of capability to provide the capacity for effective data collection and sharing. | Identifies that data be made available in a timely manner. | Identifies the importance of data integrity and consideration of duplicate entries. |
| Data Compliance Policy | Describes the commentary that jurisdictions must provide in a Statement of Assurance. | Implies the importance of data to be consistent with current standards. | Identifies a reliance on the Statement of Assurance. | Provides examples of conformity to the Australian Hospital Patient Costing Standards. | - | Identifies a reliance on the Statement of Assurance. | IHACPA is to record the time of receipt of data and evaluate compliance with timeframes. | Jurisdictions are asked to correct any errors or anomalies identified by IHACPA. |

| Policy and guidance document | Data quality dimensions | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Accuracy | Coherence | Completeness | Conformity | Coverage | Suitability | Timeliness | Uniqueness |
| Three Year Data Plan | Identifies a reliance on the Statement of Assurance. | Identifies the intention to harmonise national bodies to ensure consistency and coherence in data collection and reporting. | Identifies a reliance on the Statement of Assurance. | IHACPA to evaluate the final submission in compliance with the data request. | - | Implies suitability by stating the purpose of data. | Identifies the responsibility of jurisdictions to submit timely data.<br>Provides request and submission deadlines for Activity and NHCDC data. | Identifies the use of Unique Patient Identifiers. |
| Data Request Specifications | - | - | - | Detailed in the type and size, number of fields and valid values/notes. | Implied through the requested data items. | Implied through the requested data items. | - | Defines how Unique Identifiers are used. |

# Appendix E: List of public documents

Table 16 is a list of the publicly available documents that were accessed in this review.

### Table 16: Publicly available documents accessed

| Document ID | Documentation | Date received | Category/Description |
|---|---|---|---|
| 1 | The Addendum to the NHRA | 25 October 2024 | The Addendum to the NHRA |
| 2 | IHACPA Three Year Data Plan 2024-25 to 2026-27 | 1 November 2024 | Policy Document |
| 3 | IHACPA Data Compliance Policy 2024 | 21 October 2024 | Policy Document |
| 4 | IHACPA Data Quality Assurance Framework 2012 | 20 August 2024 | Policy Document |
| 5 | Work Program & Corporate Plan 2024-25 | 30 August 2024 | Policy Document |
| 6 | 2022-23 Data Request Specifications NHCDC Public Sector | 28 August 2024 | NHCDC Data Request Specifications |
| 7 | ABF 2021-22 Data Request Specifications:<br>• Admitted Patient Care<br>• Emergency Department Care<br>• Mental Health Care<br>• Non-Admitted Patient Care<br>• Sentinel Events<br>• Teaching, Training and Research<br>• Activity with Alternative Funding Source | 25 September 2024 | ABF Data Request Specifications |
| 8 | Independent Financial Review of the NHCDC 2020-21 Financial Year (KPMG) | 28 August 2024 | KPMG Report |
| 9 | NHCDC 2021-22 Data Quality Statements from all jurisdictions | 28 August 2024 | NHCDC Data Quality Statements |
| 10 | IHACPA National Pricing Model Technical Specifications 2024-25 | 23 September 2024 | Technical Specifications |

| Document ID | Documentation | Date received | Category/Description |
|---|---|---|---|
| 11 | IHACPA Governance framework for the development of the admitted care classifications 2022-2025 | 30 August 2024 | Background Information |
| 12 | IHACPA Pricing Framework for Australian Public Hospital Services 2024-25 | 30 August 2024 | Background Information |
| 13 | IHACPA National Efficient Cost Determination 2024-25 | 30 August 2024 | Background Information |
| 14 | IHACPA National Efficient Price Determination 2024-25 | 30 August 2024 | Background Information |
| 15 | IHACPA Understanding the NEP and NEC Determinations 2024-25 | 30 August 2024 | Background Information |

Independent Health and Aged Care Pricing Authority

Eora Nation, Level 12, 1 Oxford Street
Sydney NSW 2000

Phone 02 8215 1100
Email enquiries.ihacpa@ihacpa.gov.au

www.ihacpa.gov.au