Maddocks

Australian Bureau of Statistics

Privacy Impact Assessment

Expanded Health Data Linkage to the Person Level Integrated Data Asset

Date of analysis – 15 January 2024
Date of finalisation – 17 April 2024

This document has been prepared for the Australian Bureau of Statistics in relation to its activities.

Contents

Part A Executive Summary

1. Introduction

1.1 The Person-Level Integrated Data Asset (**PLIDA**), formerly known as the Multi-Agency Data Integration Project (**MADIP**), is a secure, person-based, research data asset that combines broad sets of information about Australian citizens, to facilitate the use and re-use of public data for research purposes.

1.2 PLIDA already includes a range of datasets about the health of Australians, including datasets which are retained on an enduring basis so that they are available for analysis by authorised researchers in accordance with the PLIDA governance arrangements. For example, PLIDA currently includes information on the usage of subsidised health care services under the Medicare Benefits Schedule (**MBS**).

1.3 The Australian Bureau of Statistics (**ABS**) is considering broader linkage of certain types of health services data described in **Attachment 2** (**Health Data**), including held by Commonwealth, State and Territory governments and not-for-profit organisations (**NFP**), to PLIDA. This **Expanded Health Data Linkage** would provide a more complete picture of the programs and services used by Australians accessing healthcare services across the country. This information would be used for research, policy development and program evaluation initiatives, with the overall aim of helping to improve the lives and health outcomes of all Australians.

1.4 The ABS has engaged Maddocks to undertake a privacy impact assessment (**PIA**) to assess the privacy impacts of the Expanded Health Data Linkage and assist ABS to take a 'privacy by design' approach in considering the appropriate framework that should be put in place to enable Expanded Health Data Linkage.

2. This PIA Process

2.1 Undertaking a PIA is consistent with the requirements of the *Privacy (Australian Government Agencies – Governance) APP Code 2017* (**APP Code**). The APP Code requires agencies to undertake a written PIA for all 'high privacy risk' projects or initiatives that involve new or changed ways of handling personal information.

2.2 Our methodology and assumptions in conducting this PIA are set out at **Attachment 1,** including descriptions of the risk ratings. A glossary of capitalised terms is at **Part F.**

2.3 This PIA:

2.3.1 considers compliance with the *Privacy Act 1988* (Cth) (**Privacy Act**), including the Australian Privacy Principles (**APPs**);

2.3.2 is intended to help ABS manage identified privacy risks and impacts of Expanded Health Data Linkages;

2.3.3 considers the safeguards that have been, or should be, put in place to secure personal information from misuse, interference or loss, or from unauthorised access, modification or disclosure; and

2.3.4 may inform the PLIDA Board and other stakeholders about Expanded Health Data Linkage.

2.1 The Australian Government is currently reviewing the Privacy Act, and while at this stage it is not clear what changes may be implemented, this PIA has taken into consideration potential reforms.

3. Summary of Findings

3.1 There can be serious negative impacts on persons if their health information is disclosed without authority, including resulting in stigma, embarrassment and discrimination.

3.2 There is a general expectation that health information, as sensitive information under the Privacy Act, should be afforded considerable protection above and beyond other personal information. Stakeholders consulted for the PIA also emphasised the need to ensure that ethical considerations (that is, issues beyond purely legal compliance) are taken into account in using Health Data.

3.3 PLIDA (previously MADIP) was established in 2015 and since that time has grown to be an invaluable data asset. PLIDA governance arrangements and processes have been strengthened since PLIDA's establishment to reflect this growth and ensure that robust measures are in place to protect personal information. A number of PIAs have been conducted in relation to PLIDA[1], including the substantive 2022 MADIP PIA Update, which have shaped current arrangements. In this PIA, we have assessed whether any further governance or other measures may need to be implemented for Expanded Health Linkage, taking into account the highly sensitive nature of health information.

3.4 In our view, the current PLIDA arrangements offer strong protections and provide a number of measures which will mean that the privacy impacts of Expanded Health Data Linkage for specific Health Data will be carefully assessed, including processes to consider the inclusion of a dataset into PLIDA, the integration of that dataset into PLIDA and use of any Health Data in approved projects under the PLIDA framework. However, we consider that the Expanded Health Data Linkage has the following heightened privacy risk given the sensitive nature of the information and potential volume of information that will be included:

3.4.1 **Key Privacy Risk:** The inherent privacy risk of re-identification[2] posed by research which uses de-identified information derived from Health Data within PLIDA, particularly where datasets contain information relating to unique patient journeys (for example, those with rare conditions that may be included in data in Cancer Registries) resulting in the potential unauthorised use and disclosure of the resulting personal information [APP 6 and 11].

3.5 We set out in **Table 1** a summary of our assessment of APP compliance for Expanded Health Data Linkage. We discuss the key APPs in greater detail at **Part D [Australian Privacy Principles Compliance]**. The recommendations set out in paragraph 4.1 and the privacy best practice recommendations set out in paragraph 4.2 are designed to address the identified risk and further enhance the privacy protections, and/or strengthen the ABS's compliance with the APPs.

Table 1: Summary of APP compliance

APP	Risk Rating	Summary
APP 1 Open and transparent management of personal information	Improvement measures to meet best practice	The PLIDA Board has made a number of commitments to transparency (see 2022 MADIP PIA Update, recommendation 1(b) and PLIDA Board Response) and we do not consider anything additional is required where these commitments continue to be implemented. However, further governance measures are recommended to assist with compliance with all APPs, in particular assurance measures that document the shared

[1] Information about PIAs conducted in relation to PLIDA can be found at: https://www.abs.gov.au/about/legislation-and-policy/privacy/privacy-impact-assessments

[2] The risk of 're-identification' is explained in greater detail at paragraph 9 of **Part B [Project Description]**.

APP	Risk Rating	Summary
		understanding between ABS and Data Custodians about the legal basis on which Health Data is integrated into PLIDA. See **Recommendation 2** and **Recommendation 4**.
APP 2 Anonymity and pseudonymity	Not relevant	Individuals will not be specifically dealing with ABS for inclusion of Health Data that relates to them for inclusion into PLIDA.
APP 3 Collection of solicited personal information	Improvement measures to meet best practice	Further measures to support the collection of Health Data into PLIDA are recommended, focussing on the public benefit of integrating a particular dataset into PLIDA. See **Recommendation 2** and **Recommendation 4**.
APP 4 Dealing with unsolicited personal information	Compliant	The requirements under APP 4 for collection of unsolicited personal information are not relevant in the context of Expanded Health Data Linkage, as all information about an individual to be included in PLIDA will be solicited by the ABS. The ABS has business-as-usual (**BAU**) processes in place should any unsolicited personal information be provided by Data Custodians as part of data integration activities.
APP 5 Notification of the collection of personal information	Compliant	Even where personal information is not technically included in PLIDA (and accordingly this is not strictly a compliance issue), the PLIDA administrative processes require Data Custodians to assure the ABS that relevant APP 5 collection notices notify individuals that information about them may be used in data integration. We do not think anything additional is required for Expanded Health Data Linkage. It is important that Data Custodians continue to work to ensure their initial data collection procedures notify individuals about how information about them may be used and disclosed.
APP 6 Use or disclosure of personal information	Compliance Risk	PLIDA processes are robust and include a range of measures to ensure data within PLIDA is used appropriately and in a way that minimises potential privacy impacts. However, given the inherently highly sensitive nature of Health Data, we consider further reasonable steps can be taken to address reidentification risks (which is a risk of disclosure of personal information). See **Recommendation 3**.
APP 7 Direct marketing	Not relevant	Not relevant, data is only used for agreed (and specified) purposes under the PLIDA governance framework, and there is no use by any organisations for direct marketing purposes.
APP 8 Cross-border disclosure	Not relevant	Not relevant, no new or changed overseas data flows will be involved in the integration of Health Data into PLIDA. There are PLIDA processes in place for consideration of any overseas researchers seeking access to PLIDA data.
APP 9 Government related identifiers	Not relevant	Not relevant, as inclusion of the expanded Health Data will not involve any new or changed use of government identifiers by any organisations.

APP	Risk Rating	Summary
APP 10 Quality of personal information	Compliant	PLIDA processes include measures to address the quality of data used in approved projects. Stakeholders did raise concerns that Health Data being transferred and stored in multiple locations may compromise the quality and security of the data. Similarly, State and Territory stakeholders highlighted the value (and reduced privacy risk) of distributed networks and coordinated national models being used to include Health Data in PLIDA, which would minimise the number of times Data Custodians are required to share the same data sets.
APP 11 Security of personal information	Compliance Risk	PLIDA processes are robust and include a range of technical and non-technical measures to ensure data within PLIDA is secure. However, linked to the risk against APP 6, the key risk for Expanded Health Data Linkage is re-identification risks, particularly for persons with unique patient journeys. We consider further reasonable steps can be taken to address reidentification risks (which is a risk of disclosure of personal information). See **Recommendation 1**.
APP 12 Access to personal information	Compliant	Once Health Data is in PLIDA, persons will not be able to access their personal information. We consider this to be reasonable given the nature of the data asset, which separates demographic information from analytical information. It is reasonable for the ABS to require that the Data Custodian should be approached if persons wish to access their Health Data, as reflected in current PLIDA BAU processes.
APP 13 Correction of personal information	Compliant	An individual will not be able to correct any personal information once Health Data is in PLIDA. It is reasonable that the Data Custodian should be approached if persons wish to correct any relevant personal information, as reflected in current PLIDA BAU processes.

4. Recommendations

4.1 This PIA makes the following **recommendation** in relation to Expanded Health Data Linkage:

Recommendation 1 Addressing Re-identification Risk	Relevant APPs
Rationale: PLIDA processes are designed so that the datasets that can be linked for approved projects contain de-identified information. However, some stakeholders have expressed the view that the risk of re-identification will increase over time for PLIDA, particularly with the inclusion of Health Data that includes unique health journeys and data of smaller jurisdictions. The PLIDA processes already include a range of technical and non-technical measures to ensure the risk of re-identification remains low. This includes the role of the ABS Disclosure Review Committee, making assessments against the Five Safes Framework, considering the risk of re-identification and recommending steps to minimise disclosure risk. However, as the negative privacy impacts of re-identification of an individual's Health Data can be potentially very significant, it is critical for ABS to be able to demonstrate that it has taken all reasonable steps to address re-identification risks associated with integrating Health Data, including as those risks change over time. **Recommendation:** We **recommend**, as a further reasonable step, that ABS build into its PLIDA arrangements a standard process for the ABS Disclosure Review Committee (or another body) to review the strategies used to address re-identification risks for PLIDA from a holistic perspective. This review or audit process could be, for example, scheduled bi-annually as part of the existing PLIDA PIA updating process and could involve the relevant ABS committee considering the robustness of current processes. In addition, there could be a documented trigger for review (such as if there was a known re-identification attempt using PLIDA data). This aims to ensure that the processes that will be employed to de-identify persons and mitigate re-identification risks for the Expanded Health Data Linkage to PLIDA remain fit-for-purpose in light of technical advances and emerging risks into the future.	**APP 6** [use and disclosure of personal information] **APP 11** [security of information]

4.2 This PIA makes the following **privacy best practice recommendations** in relation to Expanded Health Data Linkage:

Recommendation 2 Justifying inclusion of Health Data into PLIDA	Relevant APPs
Rationale: Health information is some of the most sensitive information about a person and it is generally accepted that health information should be handled with a higher degree of care. Recommendation 2(a) of the 2022 MADIP PIA Update recommended that the then MADIP Board develop and publish criteria to assess whether a new dataset should be included within MADIP (now PLIDA) and that those criteria should include consideration of the: • benefit of including the dataset within PLIDA; • the utility of including the dataset within PLIDA;	**APP 1** [open and transparent management] **APP 3** [collection of personal information]

Recommendation 2 Justifying inclusion of Health Data into PLIDA	Relevant APPs
• types of information within the dataset and restrictions on use; and • impact of including the dataset on the size and nature of the overall PLIDA asset (together, the **Minimum Criteria**) Implementing previous recommendation 2(a) (which was accepted by the MADIP Board) will be an important way to justify the expansion of PLIDA by demonstrating clearly *why* the expansion through the inclusion of new datasets containing Health Data is justified. However, depending on the timing for the full implementation of the above, the ABS should consider whether it is appropriate to take some steps in the interim to ensure that inclusion of each Health Data set into PLIDA is appropriate. **Recommendation:** We **recommend** that, while the overarching criteria for including datasets into PLIDA is being developed by the ABS and PLIDA Board, the ABS should for Expanded Health Data Linkage (potentially using the data to be received from Cancer Registries as a test case): • document the process for considering the Minimum Criteria for each dataset that contains Health Data to be included in PLIDA; • record the basis for reaching the conclusion that the public benefit factors for including a dataset outweighs the risks of including a dataset; and • continue to ensure that any data sharing agreement between the ABS and relevant Data Custodian for inclusion of a dataset containing Health Data should record the public benefit factors (see also **Recommendation 4**).	

Recommendation 3 Ethics Approval Requirements	Relevant APPs
Rationale: In a privacy context, to answer the privacy best practice question 'just because you *can* do it, *should* you do it?' requires an ethical consideration of the intended collections, use and disclosures of personal information. Maintaining ethical behaviour is also a cornerstone for the conduct of research, and the *National Statement on Ethical Conduct in Human Research* promotes ethically good human research. The National Statement may apply to approved projects under the PLIDA framework. The National Health and Medical Research Council (**NHMRC**) *Statement on consumer and community involvement in health and medical research* also recognises that appropriate consumer involvement in research should be encouraged and facilitated by researchers and research organisations because it is likely to improve the design, recruitment, conduct, and translation of research. Data ethics, more broadly, is concerned with the ethics of generating, collecting, sharing and using data, and is a burgeoning area of ethics. Work is currently underway on a whole-of-government Data Ethics Framework to 'provide guidance for the APS on best practice for ethical considerations relating to public data use and provide advice on implementation across different major use cases and agency operations'.[3]	**APP 1** [open and transparent management] **APP 6** [use and disclosure of personal information]

[3] https://www.dataanddigital.gov.au/plan/roadmap/trusted-and-secure/data-ethics-framework#:~:text=A%20whole%2Dof%2Dgovernment%20Data,use%20cases%20and%20agency%20operations.

Recommendation 3 Ethics Approval Requirements	Relevant APPs
It is critical that the use of Health Data in approved projects under the PLIDA framework meets the ethical standards for the use of data in research. Ethics approval requirements are currently part of PLIDA processes, with Data Custodians being able to stipulate ethical approval requirements and ABS tracking whether an approved project has the necessary ethics approval. **Recommendation:** We **recommend** that the ABS consider further strengthening the ethics approval arrangements for PLIDA by developing a PLIDA Ethics Framework to provide a systematic process for considering the ethical use of PLIDA datasets containing Health Data, to assist researchers and stakeholders understand the potential ethical issues. Such a framework should set out the circumstances where a Human Research Ethics Committee consideration is required as well as the circumstances where further ethics approval is not necessary. Such a Framework could also address the role of consumers or consumer advocacy groups in any approval process.	

Recommendation 4 Assurance processes in relation to the linkage of Health Data to PLIDA	Relevant APPs
Rationale: While most stakeholders generally indicated that they could see the potential benefits of Health Data being included in PLIDA, stakeholders expressed concerns that some Health Data that has been collected by States and Territories and NFPs may not be able to be used for the purposes of linking to PLIDA. Further, consideration is currently being given to whether the ABS receiving Health Data that has already been integrated outside PLIDA (for example, by the Australian Institute of Health and Welfare (**AIHW**)) should be preferred over receipt of those datasets directly from the relevant States and Territories and/or NFPs, noting that already integrated data offers administrative as well as security benefits. However, it will be important for there to be clarity about who the Data Custodian is for any such pre-integrated data. **Recommendation 4(a):** We recommend that any data sharing agreement (**DSA**), or governance arrangement, in respect of any pre-integrated data to be linked to PLIDA clearly specifies which entity (or entities) will be the Data Custodian of the relevant data under the PLIDA governance framework, and therefore responsible for assuring that the data is able to be linked to PLIDA and for approving research projects using that data. The agreed Data Custodian could be the original entity or entities who provided the data to the integrating body (such as the AIHW, which will then supply the data to the ABS for PLIDA), or the integrating body itself because of the arrangements between that body and the original supplying entity. **Recommendation 4(b):** We recommend that any DSA covering pre-integrated data set out how any data breaches will be managed, including the role of Data Custodians in the data breach management process (for example, clearing communications).	**APP 1** [open and transparent management] **APP 3** [collection of personal information]

Recommendation 4	Assurance processes in relation to the linkage of Health Data to PLIDA	Relevant APPs
Recommendation 4(c): We recommend, as a best privacy practice measure, that any DSA with a Data Custodian about the integration of a dataset containing Health Data into PLIDA includes written assurances from the Head of the Agency, or other senior personnel of the Data Custodian, such as the Chief Data Officer, that: • the Data Custodian is legally able to disclose the Health Data to the ABS for inclusion in PLIDA (e.g., because the individual consented to that disclosure at the time of collection; or because of specified legislative provisions which authorise the disclosure); and • use by the ABS of the Health Data in accordance with the DSA (if ABS complies with any restrictions set out in the DSA) will not breach any applicable legislation in the Data Custodian's jurisdiction. The basis on which these assurances are made should also be documented in the DSA (for example, that the Health Data is collected/disclosed under *XYZ Act 2024* (NSW)).		

Part B Project Description

5. Overview of Person Level Integrated Data Asset (PLIDA)

5.1 PLIDA was created as a result of a cross-portfolio government partnership of seven Commonwealth agencies: ABS, Australian Taxation Office, Department of Health and Aged Care (**DoHAC**), Department of Education, Skills, and Employment, Department of Home Affairs, Department of Social Services, and Services Australia.

5.2 The ABS is the accredited integrating authority for PLIDA and is responsible for combining the datasets in PLIDA, for providing access to PLIDA data to only those who have been authorised (**Authorised Users**) for approved research projects, for ensuring the security of data contained within PLIDA, and ensuring that all research results and other outputs from the use of PLIDA data are done in a manner that is unlikely to enable the identification of a particular person.

Benefits of PLIDA

5.3 PLIDA is a secure research data asset used by authorised researchers for approved projects in a secure, virtual access environment. PLIDA presents benefits to Australian agencies, the Australian public, researchers, and data providers (known as **Data Custodians**), including by:

5.3.1 providing Australian governments with a powerful tool for informing government decision-making;

5.3.2 supporting decisions that will help Australians live healthier, happier, and more independent lives;

5.3.3 making better use of the information that has already been collected to enhance the value of existing public data resources; and

5.3.4 making a wider range of data available for researchers from government, universities, and public policy institutes.

Governance

5.4 The PLIDA Board (comprised of the participating entities mentioned in paragraph 5.1 above[4]) is responsible for the strategic direction and oversight of PLIDA.

5.5 PLIDA is bound by the constraints of:

5.5.1 the legislation or other requirements of the Data Custodians that apply to data that they provide to PLIDA;

5.5.2 the *Census and Statistics Act 1905* (Cth) (**Census and Statistics Act**); and

5.5.3 the Privacy Act.

[4] Changes in Administrative Arrangements Orders means that the previous Department of Education, Skills, and Employment was effectively replaced by the Department of Education, and the Department of Employment and Workplace Relations.

Privacy Impact Assessments

5.6 PLIDA has evolved over time and PIAs have been conducted to assess the privacy impacts associated with the changes. Most recently, a MADIP (PLIDA) PIA update was conducted in 2022 (**2022 MADIP PIA Update**)[5] which considered:

5.6.1 new types or categories of data being included, or potentially considered for inclusion, in PLIDA, where they will be prepared for linkage in accordance with existing PLIDA processes;

5.6.2 increased volume of data (in terms of increased data from new datasets, increased variables, and increased frequency of information updates) being included, or proposed for inclusion, in PLIDA;

5.6.3 expanded outputs being permitted as part of approved PLIDA research projects; and

5.6.4 several new or changed data handling practices.

5.7 ABS has also conducted PIAs on the:

5.7.1 linkage of Cancer Institute of NSW data with PLIDA in November 2020; and

5.7.2 a supplementary consultation report in December 2022 on private sector health data, which considered the integration national intensive care data collected by the Australian and New Zealand Intensive Care Society (**ANZICS**), a not-for-profit organisation[6].

6. Overview of the Project

6.1 Data can be included in PLIDA on an enduring basis (where the dataset is retained within PLIDA and updated with new data provided by the Data Custodian over time), or data that can be included in PLIDA on a 'one-off' basis (usually for a specific project, where the data is not updated over time)

6.2 Currently, PLIDA links, as enduring linkages:

6.2.1 information on the usage of subsidised health care services under the Medicare Benefits Schedule (**MBS**) (the Data Custodian is DoHA);

6.2.2 information about the use of prescription medications and services subsidised under the Pharmaceutical Benefit Scheme (**PBS**) (the Data Custodian is DoHA);

6.2.3 information about COVID-19 and other vaccination status contained in the Australian Immunisation Register (**AIR**) (the Data Custodian is DoHA);

6.2.4 information about the health of Australians, including prevalence of long-term health condition health risk factors from the National Health Survey (**NHS**) (the Data Custodian is ABS).

The above are all Commonwealth datasets.

[5] A copy of the 2022 MADIP PIA Update can be found at: https://www.abs.gov.au/about/legislation-and-policy/privacy/privacy-impact-assessments
[6] Copies of the PIAs can be found at https://www.abs.gov.au/about/legislation-and-policy/privacy/privacy-impact-assessments

Pilot Study

6.3 One-off linkage of health services data collected by States and Territory governments has been linked to PLIDA on a project basis.

6.4 In 2020, ABS conducted a pilot on linking the following cancer patient datasets from Cancer Institute New South Wales to PLIDA (**Pilot Study**):

6.4.1 NSW Cancer Registry;

6.4.2 BreastScreen NSW; and

6.4.3 NSW Pap Test Registry.

A PIA was conducted for the Pilot Study and completed in November 2020.[7]

One-off Linkage

6.5 In 2022, ABS conducted a one-off project led by DoHAC linking the Adult Patient Database from the ANZICS which contained national intensive care data collected by ANZICS.

Benefits of Health Data Linkage

6.6 The scope of this PIA covers linkage of Health Data from Commonwealth and State and Territory governments and NFPs[8].

6.7 **Health Data**[9] has a broad meaning and includes the following types:

6.7.1 administrative health data,

6.7.2 patient health and disease data (including outcomes of medical tests, biomedical markers and genetic testing for diagnosis of diseases or conditions),

6.7.3 health survey data,

6.7.4 ambulance and patient transport data,

6.7.5 emergency department and outpatient data, and

6.7.6 hospital data.

6.8 Health Data types are not mutually exclusive, and data may be categorised across more than one Health Data type. A description of each type and examples of data items are provided in **Attachment 1**.

6.9 Specific Health Data out of scope for this PIA includes: My Health Record data, private health insurance data, clinical trial data and genetic data (except outcomes of medical tests[10], biomedical markers and genetic testing for diagnosis of diseases or conditions).

6.10 To inform the privacy considerations of the Expanded Health Data Linkage and the development of a consistent framework to support any Expanded Health Data Linkage, the ABS is considering the datasets maintained, or potentially to be provided, from the following entities as case studies:

[7] A copy of the PIA for the Pilot Study can be found at: https://www.abs.gov.au/about/legislation-and-policy/privacy/privacy-impact-assessments
[8] Organisations in control of the data.
[9] This is based on the definition of 'health information' in s 6FA and 'health service' in s 6FB of the Privacy Act.
[10] The ABS has indicated that it intends to use a general definition of 'medical tests' to mean tests that are designed to diagnose, treat or monitor the health of individuals.

6.10.1 Cancer Institute New South Wales;

6.10.2 Cancer Alliance Queensland; and

6.10.3 Cancer Council Victoria,

(together, **Cancer Registries**).

6.11 The Cancer NSW datasets are:

6.11.1 hospital admissions and emergency department attendance in relation to cancer patients attending NSW hospitals; and

6.11.2 updated existing datasets from the Pilot Study on an annual basis.

6.12 The Cancer QLD dataset is the Queensland Oncology Repository, which includes details on:

6.12.1 investigations (such as, pathology and radiology) and results of investigations;

6.12.2 diagnoses;

6.12.3 cancer treatments of individuals;

6.12.4 hospitalisations of cancer patients from Queensland health data; and

6.12.5 cancer treatment outcomes.

6.13 The Cancer Victoria dataset is the Victorian Cancer Registry, which includes a registry of Victorian residents diagnosed with cancer between 1982 –2023. .

PLIDA linkage infrastructure

6.14 PLIDA contains high-value, person-centred and regularly updated datasets that aim to comprehensively cover the Australian population. The data contained within PLIDA is currently primarily made up of Commonwealth data provided by Data Custodians who are Australian Government entities (or other entities authorised by Data Custodians to provide the data), that is linked by the ABS for PLIDA. If individuals in those datasets are reasonably identifiable, the datasets will be considered to include personal information (including sensitive information (as defined in the Privacy Act)).

6.15 ABS has built a central linkage infrastructure within PLIDA. It does this by linking the datasets supplied by Data Custodians to a central linkage infrastructure called the 'Person Linkage Spine' (**Spine**). The Spine creates the capacity for separate source databases to be linked when required for research projects, for use by Authorised Users for approved projects.

6.16 PLIDA adheres to the 'separation principle' and 'functional separation'[11] in receiving, storing and curating data for all integration projects. This means that for each dataset, data that contains identifying information about each individual that is the subject of the dataset ('linkage information'), is stored and handled separately from other data about that individual ('analytical information'). Access to these different information sets is restricted so that there is no access to both sets of information simultaneously. Functional separation means that ABS staff members undertaking data linkage only have access to the information they need to perform their assigned roles.

[11] Further details on the separation principle and functional separation can be found at: https://www.abs.gov.au/about/data-services/data-integration/keeping-integrated-data-safe.

6.17 ABS is responsible for ensuring that the PLIDA data that it makes available to authorised researchers in the secure DataLab is provided in a manner that is not likely to enable the identification of an individual (and therefore meets the requirements to be 'de-identified' under the Privacy Act and the Census and Statistics Act).

6.18 ABS currently has a range of security arrangements in place for the IT systems used for PLIDA to protect PLIDA data, which:

6.18.1 conform with security arrangements set out in the Australian Government Information Security Manual (**ISM**);

6.18.2 ensure that data collection, linkage and assembly activities for PLIDA datasets are only conducted by a dedicated team in the Secure Data Integration Environment (**SDIE**);

6.18.3 includes a secured internet gateway which is reviewed annually by the Australian Signals Directorate (**ASD**); and

6.18.4 includes an ongoing program of security audits and system accreditations, including the Information Security Registered Assessors Program (**IRAP**).

6.19 Further, PLIDA datasets are handled in accordance with a range of additional privacy protection practices, including:

6.19.1 the Five Safes Framework is applied to ensure access to PLIDA data is appropriate. The Five Safes Framework is designed to facilitate safe data release using five elements (Safe People, Safe Projects, Safe Settings, Safe Data, and Safe Outputs) which are assessed independently, but also considered as a whole for each instance of data access;[12]

6.19.2 data may also be added to PLIDA via a once-off linkage (for a specific research project or projects and will not be retained following the completion of the project(s), or as a part of an enduring analytical asset that is separate to the Spine; and

6.19.3 the ABS is transparent about linkages with PLIDA and approved projects that make use of PLIDA data, through registers and other information published on the ABS website.

[12] Further details on the Five Safes Framework can be found at: https://www.abs.gov.au/about/data-services/data-confidentiality-guide/five-safes-framework.

Part C Key Concepts

7. Analysis of Information Handled as Part of Health Data Linkage.

7.1 In this Part, we set out information about some concepts that are generally relevant to our analysis of Expanded Health Data Linkage against the APPs (discussed in **Part D [Australian Privacy Principles Compliance]**).

8. De-identified information

8.1 A significant threshold question is the extent to which elements of Expanded Health Data Linkage involve the handling of 'personal information'. The Privacy Act only covers information that is personal information.

8.2 Section 6(1) of the *Privacy Act* defines 'personal information' as:

> ***personal information*** means information or an opinion about an identified individual, or an individual who is reasonably identifiable:
>
> (a) whether the information or opinion is true or not; and
>
> (b) whether the information or opinion is recorded in a material form or not.

8.3 PLIDA relies on the use of the Person Linkage Spine (**Spine**). The creation of the Spine relies on the use of personal information of individuals. The Spine is a tool that contains direct identifiers for each individual in a dataset within PLIDA (such as their names, dates of birth, addresses) and maps this to each reference number for that person in each of the datasets stored within PLIDA.

8.4 On the other hand, the analysis to be undertaken as part of approved projects under the PLIDA framework relies on the use of content data (which is analytical data without any direct identifiers of the data subjects, such as their names, dates of birth or contact information). As articulated by the Office of the Australian Information Commissioner (**OAIC**), de-identified information is information which has undergone a process of de-identification so that a person is no longer 'reasonably identifiable'. We consider de-identifying information to be a privacy enhancing measure, as it mitigates against the risk of unauthorised disclosure of, or access to, personal information.

8.5 The PLIDA governance arrangements ensure best practice data linkage approaches are adopted. This includes minimising data movements, restricting access to personal information within the ABS, and application of the separation principle, which ensures that personal information used for linkage is handled separately to content data. In addition, at the project stage, the PLIDA governance arrangements include additional steps to ensure that the risk of re-identification is sufficiently low (see paragraph 9 below). We consider these measures to be in line with current best practice to de-identifying personal information.

8.6 These measures mean that Health Data used for linkage is personal information, but content Health Data which is to be incorporated into PLIDA can be treated as de-identified in accordance with current PLIDA practices. Where information is de-identified it is not 'personal information' and therefore is not subject to the Privacy Act. However, from a policy perspective, we consider there is merit in applying the APPs as good governing principles for all data handled as part of PLIDA and our analysis proceeds on this basis.

9. Re-identification risks

9.1 In any project that involves de-identifying personal information or use of de-identified information, consideration also needs to be given to the potential risk of re-identification, which will then mean the information is personal information within the scope of the Privacy Act.

9.2 The Office of the Australian Information Commissioner (**OAIC**) guide *'De-identification and the Privacy Act'* states that it is not necessary to remove the risk of re-identification entirely. Rather, controls and treatments should be implemented by those sharing or releasing data to mitigate the risk until it is 'very low' and there is no reasonable likelihood of re-identification of occurring.

9.3 Addressing re-identification risks is an important aspect of ABS's functions. The Census and Statistics Act, which governs the collection, use and disclosure of data for research and statistical purposes, provides that the ABS cannot release information of a personal or domestic nature in a manner that is likely to enable the identification of that person.

9.4 PLIDA relies on a number of technical and non-technical measures and controls to address the potential risk of re-identification, including:

9.4.1 the implementation of the separation principle, to ensure that no-one will have access to both identifying information and content data about the same individual at the same time;

9.4.2 treatment of content data to minimise risk of re-identification (for example, data above or below an upper boundary is not displayed as an actual amount, but simply as being 'above' or 'below' the boundary);

9.4.3 compliance with data sharing agreements which (amongst other things) which set out the strict purposes for which data can be used and include express restrictions against attempts to re-identify any individual from their content data;

9.4.4 as part of the Five Safes Framework assessment, the ABS Disclosure Review Committee considers the risk of re-identification and recommends steps to minimise disclosure risk;

9.4.5 training of all people involved with the data around their responsibilities to keep data safe;

9.4.6 monitoring and auditing of activities involving use of the data;

9.4.7 strict controls over the outputs that may be generated from approved research projects;

9.4.8 robust security and access protections for the Spine and National Linkage Map, given the potential for those underlying infrastructure components to be used for re-identification purposes; and

9.4.9 the deterrent effect of the criminal penalties that researchers are liable for if they misuse the data, such as seeking to identify individuals. All researchers accessing data through the DataLab are required to sign undertakings (which are binding) that they will only use data for statistical or research purposes, and they will make no attempt to use the data with any other data.

9.5 While these measures arguably already address the re-identification risk posed in the current environment, it is important to appreciate the potential size and scale of Expanded Health Data Linkage, and that it involves exchange of information between different entities and use of that information by those entities.

9.6 If the persons to whom de-identified information relates were to be re-identified, it could have adverse impacts for those persons if that information was to be disclosed or misused, including embarrassment, psychological harm, or other consequences. Persons with rare conditions, or with unique patient journeys, are more susceptible to re-identification.

9.7 This raises the Key Privacy Risk for Expanded Health Data Linkage:

9.7.1 **Key Privacy Risk**: The inherent privacy risk of re-identification posed by linking of de-identified information derived from Health Data within PLIDA, particularly where datasets contain information relating to unique patient journeys (for example, those with rare conditions that may be included in data in Cancer Registries) resulting in the potential unauthorised use and disclosure of the resulting personal information [APP 6 and 11].

9.8 Our **Recommendation 1** is intended to build on current PLIDA arrangements by building in review mechanisms to further address the potentially increased re-identification risk posed by Expanded Health Data Linkage.

10. Ethics and Privacy Laws

10.1 Stakeholders during the consultation for the PIA emphasised the importance of ethics approval as another means to reduce risk in relation to research projects. Some stakeholders were concerned that currently not all data projects require an ethics assessment and some, particularly non-academic projects, may be able to bypass any form of ethics approval.

10.2 In Australia, the *National Statement on Ethical Conduct in Human Research 2023* (**National Statement**) has been developed as a set of guiding standards for individuals, institutions and organisations conducting human research. It also provides guidance to review bodies, including Human Research Ethics Committees (**HRECs**) in relation to conducting ethical review of research.

10.3 Two of the guiding principles in the National Statement, are that:

10.3.1 the proposed research is justifiable by its potential benefit which may include 'contribution to knowledge and understanding, to improved social welfare and individual wellbeing, and to the skill or expertise of researchers',[13] and

10.3.2 that the likely benefit of the research must justify any risks of harm or discomfort to participants.[14]

10.4 Whether ethics approval is required is dependent on any applicable statutory obligations and funding requirements of particular research projects, and is otherwise a matter of best practice. For example, s 12AD(c) of the *Therapeutic Goods Regulation 1990 (Cth)* mandates that the National Statement is to be applied in relation to the use of therapeutic goods in a clinical trial and for experimental purposes involving humans.

10.5 It has also been recognised that appropriate consumer involvement in research should be encouraged and facilitated by researchers and research organisations because it is likely to improve the design, recruitment, conduct, and translation of research (see NHMRC's *Statement on consumer and community involvement in health and medical research*).

10.6 Currently, whether ethics approval is required for research projects seeking to use PLIDA data is dependent on the requirements of the Data Custodian. For research projects requiring Aboriginal or Torres Strait Islander data, the ABS' standard practice is for the ABS Centre for Aboriginal and Torres Strait Islander Statistics to undertake a cultural safety assessment, focussing on a proposed project's interaction with First Nations peoples, the risk of deficit narrative or harm to communities.

[13] National Statement on Ethical Conduct in Human Research Principle 1.1(a), 10.
[14] As above, Principle 1.6, 10.

10.7 The Privacy Act deals with ethics approval in a limited context. An APP entity seeking to collect sensitive information for research, or to use or disclose personal or sensitive information which has already been collected for the secondary purpose of research, must either obtain the consent of the relevant individual or rely on one of the research exceptions in the Act. The Act currently has separate research exceptions for agencies and organisations. For example, section 95 permits agencies to deviate from the APPs in the course of medical research, if done in compliance with guidelines issued by the NHMRC made for the purposes of that section.[15]

10.8 However, the exceptions in the Privacy Act are only relevant to the handling of 'personal information'. The PLIDA arrangements, particularly the application of the separation principle which separates demographic information from analytical information, are designed to ensure that datasets integrated into PLIDA and used for an approved project involve the handling of de-identified information. Further, any resulting outputs from approved projects are checked as a measure to address any potential re-identification risks.

10.9 We also note that currently a burgeoning area of ethics development is 'data ethics' more broadly, which is concerned with the ethics of generating, collecting, sharing and using data. Work is currently underway on a whole-of-government Data Ethics Framework to '*provide guidance for the APS on best practice for ethical considerations relating to public data use and provide advice on implementation across different major use cases and agency operations*'.[16]

10.10 The above discussion demonstrates the complex interplay between ethics and legal compliance. Irrespective of the strict legal requirements, in the context of Expanded Health Data Linkage, involving the potentially significant increase in the handling of some of the most sensitive information about people, we consider it critical from a privacy best practice perspective that ABS and Data Custodians can answer the privacy best practice question 'just because you *can* do it, *should* you do it?'. This requires an ethical consideration of the intended collections, use and disclosures of personal information.

10.11 We therefore consider that the PLIDA arrangements for Expanded Health Data Linkage can be further strengthened by implementing:

10.11.1 **Recommendation 2** – so that the ABS and Data Custodians can justify the inclusion of datasets containing Health Data into PLIDA; and

10.11.2 **Recommendation 3** – a development of a PLIDA Ethics Framework will assist ABS, Data Custodians and researchers ensure that approved projects meet the ethical standards for the use of data in research.

[15] Privacy Act s 95.
[16] https://www.dataanddigital.gov.au/plan/roadmap/trusted-and-secure/data-ethics-framework#:~:text=A%20whole%2Dof%2Dgovernment%20Data,use%20cases%20and%20agency%20operations.

Part D Australian Privacy Principles Compliance

11. Overview

11.1 In this Part we discuss the following key privacy issues (set out in a series of tables below) which we consider to be the most relevant to Expanded Health Data Linkage, and consider whether any further steps are required in addition to those already in place for PLIDA:

11.1.1 governance and oversight (as reflected in APP 1);

11.1.2 openness and transparency (as reflected in APP 1);

11.1.3 collection of Health Data for inclusion in PLIDA (as reflected in APP 3 and APP 5);

11.1.4 use of Health Data in PLIDA (as reflected in APP 6);

11.1.5 disclosure of information (as reflected in APP 6); and

11.1.1 data security (as reflected in APP 11.1).

11.2 Our analysis does not address those elements of the APPs which reflect the ABS's broader compliance obligations.

11.3 A summary of our analysis against all the APPs, which includes a risk rating based on the descriptors at **Attachment 1**, is set out in **Table 1** in **Part A [Executive Summary]**.

12. Openness and Transparency

Issue	Discussion of issue	Potential further steps
Principles behind APP 1 - Open and transparent management of personal information APP 1 is intended to ensure that entities manage personal information in an open and transparent way. It is also intended to ensure that individuals dealing with that entity are provided with information about how the entity manages personal information.		
12.1 **Transparency of data collection and use**	Failure to adequately communicate the purposes of Expanded Health Data Linkage, and the privacy safeguards that are built (or will be built) into its processes and solutions, is likely to undermine the confidence of the community, Data Custodians, privacy regulators, Ministers, and the public, as well as potentially challenge the purpose and desired outcomes of PLIDA. The ABS website already includes a range of information about PLIDA that assists in meeting this principle. The PLIDA Board has also made a number of commitments to transparency. In response to 2022 recommendation 1(b) to the 2022 MADIP PIA Update, the PLIDA committed to: *The MADIP Board and the ABS remain committed to transparency and engagement with stakeholders. The ABS will continue to refine and enhance online materials to increase transparency for MADIP. Where possible, advance notification of the consideration of new categories of data or new data handling practices in MADIP as well as reasons for including/not including data in MADIP will be provided in on-line materials.*	Continuing to action previous commitments made by the MADIP (now PLIDA) Board in relation to transparency supports Expanded Health Data Linkage. We do not consider, from a compliance or best privacy perspective, that ABS or the PLIDA Board need to specifically do anything additional to promote transparency of the data collection and use associated with Expanded Health Data Linkage. However, as is the intention behind APP 1, ABS and the PLIDA Board should continue to look for opportunities to be as transparent as possible about PLIDA's operations. We also make the observation that an important aspect of transparency for PLIDA and the integration of Health Data is what people are told at the time of collection by the relevant Data Custodian of the Health Data. As set out in recommendation R2 of 2019 MADIP PIA Update, it is important for Data Custodians to be 'responsible for collection notices to enhance transparency about their disclosure of personal information to the ABS for MADIP by taking reasonable steps to update notices'.

13. Governance and oversight arrangements

Issue	Discussion of issue	Potential further steps
Principles behind APP 1 - Open and transparent management of personal information APP 1 also requires entities to take reasonable steps to implement practices, procedures and systems that will ensure compliance with the APPs.		
13.1 **Appropriate data governance arrangements**	Good data governance arrangements will assist ABS to meet all APPs (in particular APP 1 and APP 6). Having robust governance arrangements and processes between the ABS and Data Custodians will demonstrate how risks will be mitigated and issues addressed in Expanded Health Linkage. The processes under the PLIDA governance framework, shaped over time as the asset has grown, already include a number of robust governance processes, including the requirement for data sharing agreements to be entered into between Data Custodians and ABS to govern any data provided to ABS, the oversight role of the PLIDA Board, and assurance steps taken by the ABS to review proposed outputs of approved projects. Details of the operation of PLIDA can be found at: https://www.abs.gov.au/about/data-services/data-integration/integrated-data/person-level-integrated-data-asset-plida	Noting that continuous improvements can always be made to how entities handle personal (and other) information, we have made the following privacy best practice recommendations that support compliance with the other APPs (and are discussed in the tables below in more detail): **Recommendation 2**, **Recommendation 3**, **Recommendation 4**

14. Collection of Health Data for Expanded Health Data Linkage

Issue	Discussion of issue	Potential further steps
Principles behind APP 3 - Collection of solicited personal information **Principles behind APP 5 - Notification of the collection of personal information** APP 3 is intended to minimise the collection of personal information to that which is reasonably needed by the relevant entity to undertake its particular functions and activities. This is sometimes referred to as the 'data minimisation principle'. Further limitations should apply to the collection of sensitive information (such as prohibiting collection unless the relevant individual has consented to that collection, where the collection is authorised or required by another law, or other specific circumstances apply where it is reasonable to collect that sensitive information). APP 5 requires an entity that collects personal information about an individual to take reasonable steps to notify the individual of certain matters (referred to as "APP 5 matters"), or otherwise ensure that the individual is aware of those matters.		
14.1 **Data needs to be legally able to be collected**	Health Data to be included in PLIDA will have been initially collected by Data Custodians for performance for their particular functions or activities. Health Data needs to be able to be disclosed by the Data Custodian to ABS for PLIDA (and then collected by ABS). Expanded Health Data Linkage does not involve collection of any new information directly from any individual. Entities should not seek to 'over-collect' personal information which it might be 'nice to have' but cannot be justified as reasonably required. This APP needs to be met at the time of collection of personal information. For sensitive information (such as health information), APP 3.3 requires that the individual must have consented to the collection, as well as that the collection must be reasonably necessary for, or directly related to, an entity's functions and activities. Alternatively, an exception in APP 3.4 needs to apply (such as that the collection is authorised by law). ABS undertakes extensive liaison with Data Custodians, including obtaining assurances that the Data Custodian is legally able to provide the relevant data for inclusion in	In our view, while it is predominantly a matter for the relevant Data Custodian (i.e. the State, Territory or NFP) about whether any Health Data intended to be integrated into PLIDA is subject to any constraints under their respective regimes, we consider it is important for ABS to be able to demonstrate it has undertaken proper due diligence in relation to its collection of Health Data that is proposed to be linked. **Recommendation 4** sets out an assurance processes regarding the data sharing agreement that is to underpin the data linkage of any dataset that contains Health Data. Recording assurances from Data Custodians in a data sharing agreement signed by a relatively senior official from the Data Custodian would provide a level of assurance to not only ABS, but also stakeholders. We think this is appropriate given the sensitivity of Health Data. Further, as discussed in paragraph 10 in **Part C [Key Concepts],** in addition to considering the legal requirements for including a dataset within PLIDA, we consider that it is critical as privacy best practice for ABS and Data Custodians to be able to justify why it is in the public benefit to include a dataset containing Health Data into PLIDA (see **Recommendation 2**).

Issue	Discussion of issue	Potential further steps
	PLIDA, and whether there are any limitations on use of the content data for research.	
14.2 **APP 5 collection notices**	APP 5 requires an entity that collects personal information about an individual to take reasonable steps to notify the individual of certain matters (referred to as "APP 5 matters"), or otherwise ensure that the individual is aware of those matters. Many States, Territories and NFPs collecting data to be linked to PLIDA will have limited opportunity to provide an APP 5 notice after collection. Data Custodians are responsible for ensuring relevant notifications are in place in line with applicable privacy laws in their jurisdiction. The PLIDA administrative processes require Data Custodians to assure the ABS that relevant APP 5 collection notices notify individuals that information about them may be used in data integration.	We do not think anything additional is required for Expanded Health Data Linkage. It is important that Data Custodians continue to work to ensure their initial data collection procedures notify individuals about how information about them may be used and disclosed.
14.3 **Consent**	The majority of Health Data that is likely to be used for Expanded Health Data Linkage will have been initially collected from individuals under a wide range of circumstances and may be protected by a wide range of legislation (including privacy legislation). The data will also need to be collected by ABS for inclusion in PLIDA. Under APP 3, the ABS does not need the consent of individuals to collect personal information where it is reasonably necessary for, or directly related to, ABS's functions or activities (such as PLIDA), unless that personal information is also sensitive information (such as health information). In that case, consent is necessary, or a relevant exception needs to apply. An example of a relevant exception is where collection of sensitive information is required by law.	We do not consider that any further specific measures need to be taken to address consent, in our view implementing the privacy best practice recommendations in **Recommendation 2**, **Recommendation 3** and **Recommendation 4** will support the ABS and Data Custodians in explaining the role of consent for Expanded Health Data Linkage.

Issue	Discussion of issue	Potential further steps
	For some datasets, the collection of 'sensitive information' may also be facilitated through sections 95 and/or 95A of the Privacy Act, so that personal information (including health information) can be collected, used and/or disclosed in accordance with strict guidelines for the conduct of medical research, or for research or the compilation of analysis of statistics relevant to public health. Stakeholders expressed concern about the unspoken 'social licence' in PLIDA's data governance arrangements. Stakeholders questioned whether adding Health Data to PLIDA is moving too far away from the Data Custodian's original purpose for data collection of that data, and whether the necessary consents have been obtained. As discussed above, this is predominantly an issue that needs to be addressed at the time of collection by Data Custodians. We also note that consent may not be legally necessary where de-identified information is used (as will be the case for content data included in PLIDA), however, the commitments to transparency, as described in item 12.1 above, will also assist Data Custodians and the ABS in communicating the objectives of PLIDA, including providing an opportunity to explain to the public the role of consent in the context of PLIDA.	

15. Use and disclosure of linked data from Health Data Linkage

Issue	Discussion of issue	Potential further steps
Principles behind APP 6 - Use and disclosure personal information APP 6 is intended to restrict personal information that was collected for one purpose (the primary purpose) from being used or disclosed for another purpose (a secondary purpose), except in specific circumstances (including where the individual has consented to that secondary use or disclosure, or where the use or disclosure is required or authorised by another law).		
15.1 **Uses of Health Data for PLIDA research projects are in line with community expectations**	It is important that PLIDA maintains its social licence. Part of building this social licence is understanding and responding to community concerns and meeting community expectations. ABS currently consults with a range of stakeholders in formulating PLIDA arrangements (including consultation as part of this PIA) to gauge community expectations. We note in particular the role of specialists teams in ABS that consider the cultural safety aspects of data relating to Aboriginal and Torres Strait Islander peoples.	In our view, the nature of Health Data, being some of the most sensitive information about individuals, is such that it warrants closer scrutiny about how it is used. In addition to ensuring there is a strong public benefit for inclusion of any particular dataset in PLIDA (**Recommendation 2**), it will be critical for ABS to continue to build social licence about use of that data in PLIDA. In our view, the PLIDA arrangements can be strengthened by the development of an ethical framework, which goes beyond legal compliance consideration to support the ethical use of Health Data in PLIDA (**Recommendation 3**).
15.2 **Appropriate governance and arrangements for acceptable data uses**	In our view, the PLIDA governance framework and processes are robust, with arrangements that clearly set out approval processes for the use of data in PLIDA (including requirements for the Data Custodians of the data in PLIDA to provide approval for all integration and use of that data for any research projects). We consider that these should be sufficient if applied to Health Data within PLIDA.	
15.3 **Appropriate controls for access to data**	We note that the 2022 PLIDA PIA considered the requirements to be placed on those seeking access to the DataLab (researchers). Again, we consider that the current controls should be sufficient if applied to Health Data within PLIDA.	

Issue	Discussion of issue	Potential further steps
15.4 **Appropriate controls for outputs**	In our view, the current controls are robust, and should be sufficient if applied to Health Data within PLIDA, to ensure data released is de-identified and in aggregate form. We observe that there may be benefits from a social licence perspective if research outcomes are required to be made public and available from the PLIDA website, however that is a matter that is best considered on a case-by-case basis by at the project approval stage in accordance with the current governance arrangements.	

16. Data security

Issue	Discussion of issue	Potential further steps
Principles behind APP 11.1 - Security of personal information APP 11.1 is intended to ensure that entities take such steps as are reasonable to protect personal information from misuse, interference, and loss, and from unauthorised access, modification, or disclosure; and that they take reasonable steps to destroy the information or to destroy or de-identify personal information that they no longer need.		
16.1 **Data held securely**	Given the sensitivity of the content (analytical) Health Data, and the potential privacy impacts if that data was to be inappropriately accessed or disclosed and individuals were able to be re-identified, it is critical for it to be secured. The implementation of the 'separation principle' as explained in paragraph 6.16 in **Part B [Project Description]** is an important mitigation strategy for PLIDA to reduce the risk of identification in relation to content data. PLIDA also has a range of strong technical security arrangements (see paragraphs 6.18 and 9.3 in **Part B [Project Description]**) which protects the data within PLIDA.	In our view, the security arrangement for PLIDA are robust, however, as discussed in paragraph 9 of **Part C [Key Concepts]**, the key privacy risk for Expanded Health Data Linkage is the potential for an increased risk of re-identification, with associated potentially serious adverse consequences for individuals if re-identification was to occur. Given this key risk, it is important for ABS to demonstrate that it has taken all reasonable steps to protect Health Data integrated into PLIDA. Our **Recommendation 1** is intended to build on current PLIDA arrangements by building in review mechanisms to further address the re-identification risk posed by Expanded Health Data Linkage. Stakeholders also highlighted the value of storing data centrally to avoid producing copies of the data. There were concerns that Health Data transferred to multiple locations may compromise the quality and security of the data. Similarly, State and Territory stakeholders highlighted the value (and reduced privacy risk) of distributed networks and coordinated national models being used to include Health Data in PLIDA which would minimise the number of times that Data Custodians are required to share the same data sets. We support the proposal to, where possible, use pre-integrated Health Data sets that are already maintained by other integrating authorities (for example, the AIHW). We consider this to be privacy enhancing feature because using personal information that has already been transformed into analytical data sets will minimise handling risks and reduce risks associated with the further transfer

Issue	Discussion of issue	Potential further steps
Principles behind APP 11.1 - Security of personal information APP 11.1 is intended to ensure that entities take such steps as are reasonable to protect personal information from misuse, interference, and loss, and from unauthorised access, modification, or disclosure; and that they take reasonable steps to destroy the information or to destroy or de-identify personal information that they no longer need.		
		of additional copies of the same data by the original data provider. We note that if such arrangements are used, it will be important to specify which entity will be the Data Custodian under the PLIDA governance arrangements, including being responsible for approving research projects (for example, whether the original data provider will be the Data Custodian, or the AIHW will be the Data Custodian and provide approval in accordance with its own arrangements with its data provider(s)). See **Recommendation 4(a)**.
16.2 **Data breach management**	Some Data Custodians will be subject to State or Territory data breach management obligations or voluntary schemes, so there will need to be clarity about whether the Privacy Act regime and/or other regimes will apply to data stored in PLIDA, and how these can be best implemented to ensure compliance.	To assist in the management of any potential data breaches, we recommend that the arrangements between the Data Custodian and ABS are clearly documented. See **Recommendation 4(b)**.

Part E Glossary

Definitions	
ABS	means the Australian Bureau of Statistics.
AIHW	means the Australian Institute of Health and Welfare.
ANZICS	means the Australian and New Zealand Intensive Care Society.
APP Code	means the *Privacy (Australian Government Agencies – Governance) APP Code 2017*.
APP Guidelines	means the OAIC's *Australian Privacy Principles guidelines.*
APP, or Australian Privacy Principle	has the meaning given to it in the Privacy Act.
ASD	means the Australian Signals Directorate.
Authorised User	means a person authorised to access PLIDA data.
Cancer Registries	refers to datasets of the Cancer Institute New South Wales, Cancer Alliance Queensland and Cancer Council Victoria
Census and Statistics	means the *Census and Statistics Act 1905* (Cth).
Data Custodian	means the entity which is allocated responsibility for use of a dataset under the PLIDA governance framework. A Data Custodian is usually the entity which provides the data to the ABS but could be another entity nominated in the relevant data sharing agreement as the entity having responsibility over the data set.
DSA	means data sharing agreement which governs the provision of a dataset from one entity to another.
Expanded Health Data Linkage	means the broader linkage of additional Health Data to PLIDA on an enduring basis, as described in this PIA.
Health Data	means data about health and health services, collected and held by Commonwealth and State and Territory governments and not-for-profit organisations, as further described in **Attachment 1**.
HREC	means a Human Research Ethics Committee.
IRAP	means the Information Security Registered Assessors Program.
ISM	means the Information Security Manual.
MADIP	means the Multi-Agency Data Integration Project (the previous name of PLIDA).
MBS	means the Medicare Benefits Schedule.
National Statement	means the National Statement on Ethical Conduct in Human Research 2023.

Definitions	
NFP	means a not-for-profit organisation.
NHS	means the National Health Survey.
OAIC	means the Office of the Australian Information Commissioner.
PBS	means the Pharmaceutical Benefit Scheme.
personal information	has the meaning given in section 6 of the Privacy Act.
PIA	means this privacy impact assessment.
PLIDA	means the Person-Level Integrated Data Asset (formally MADIP) which is a secure, person-based, research data asset that combines broad sets of information about Australian citizens, to facilitate the use and re-use of public data for research purposes.
Privacy Act	means the *Privacy Act 1988* (Cth).
SDIE	means the ABS Secure Data Integration Environment.
sensitive information	has the meaning given in section 6 of the Privacy Act.
Spine	means the Person Linkage Spine maintained by ABS.

Attachment 1 Methodology and Assumptions

17. Our Methodology

17.1 This PIA has been conducted in accordance with the *Guide to undertaking privacy impact assessments* issued by the OAIC, using the methodology in the table below.

Stage	Description of steps
1.	**Plan for the PIA:** To assist in ensuring that we had correctly understood the background and to increase awareness of the issues we considered were likely to be important for this PIA, we prepared an Issues Guidance document to assist in the design of the relevant stakeholder engagement process.
2.	**Project Description**: We prepared an initial draft Project Description, which described our understanding of the Project. This draft was refined and then finalised following feedback from ABS.
3.	**Stakeholder consultation:** Undertaking consultation with stakeholders was an essential part of conducting this PIA. It provided both an opportunity to inform stakeholders about the Expanded Health Data Linkage, and to listen to stakeholder views on the proposed privacy management arrangements. We assisted ABS to prepare and plan the stakeholder consultation workshops, including advising on stakeholder consultation lists and commenting on draft material for the consultation. We produced a Consultation Report, which summarises the feedback. **Attachment 2** sets out the list of stakeholder organisations that participated in the consultation.
4.	**Privacy impact analysis and compliance check**: In this step we focussed on identifying the privacy impacts of each change, and its compliance against the relevant APPs and privacy best practice. In undertaking our analysis, we considered and applied the Australian Privacy Principles Guidelines (**APP Guidelines**) issued by the OAIC, which outline the mandatory requirements of the APPs, how the OAIC will interpret the APPs, and matters that may be taken into account when assessing compliance with the *Privacy Act*. In addition to the valuable insights gained through the stakeholder consultation process, we also used our knowledge of Australian community expectations gained from research and related work in identifying and analysing privacy risks. Where we identified a privacy risk, we allocated a rating that reflects our assessment of the compliance risk against the relevant APP(s), based on the following descriptors:

Compliance Rating	Description of Compliance Rating
Significant further work required	*This rating indicates that we consider that extensive work is required before the Commonwealth Partners will be able to be fully satisfied about compliance with the relevant APP and best privacy practice.*
Compliance risk	*This rating indicates that we have identified issues which we consider should be addressed in order to further enhance the privacy protections for individuals.*
Improvements to meet best practice	*This rating indicates that we have identified measures that can be taken by the Commonwealth Partners to meet privacy best practices.*

Stage	Description of steps		
		Compliant	*This rating indicates that the APP is either not relevant in the circumstances or that we have concluded that no further mitigation steps are required.*
5.	**Privacy management and addressing risks:** We considered potential mitigation strategies that could reduce or remove the privacy impacts and risks identified during the previous step, and developed our recommendations.		
6.	**Draft report:** We prepared a draft version of this PIA report for review by ABS.		
7.	Following review of the draft report by ABS, we further refined our analysis and potential mitigation strategies as required to ensure that privacy risks were appropriately considered and addressed.		
8.	We finalised this PIA report, including incorporation of ABS responses to our recommendations.		

18. Assumptions and Qualifications

18.1 We have conducted our analysis on the basis that the factual information provided ABS (as set out in **Part B [Project Description and Information Flows**] is up-to-date, correct, and complete.

18.2 This PIA only considers the privacy implications of the Project as described in this PIA report. We understand that any material changes to the implementation of PLIDA that impact the handling of personal information will be subject to an update to this PIA, or other privacy advice, as required.

18.3 Our PIA has been undertaken from the perspective of ABS and not any other entity.

18.4 Our analysis is based upon the provisions of the Privacy Act, and associated case law and guidance material, as at the date of analysis on the cover page of this PIA report. While we have endeavoured to take into account relevant proposed reforms of the Privacy Act discussed in the Privacy Act Review Report released by the Attorney General's Department, and the Australian Government's responses, the ABS will need to separately consider the impact on the Project of any proposed legislation to effect any reforms, if and when it is introduced.

Attachment 2 Health Data Types and Descriptions

Broad types of health data	Description and data items

1. Administrative health data	Includes information that is routinely collected by governments and non-government organisations as part of administering the delivery of health related services.
	
	Data items may include:
	• patient demographic information
	• Medicare records (e.g., Medicare Benefits Schedule and Pharmaceutical Benefits Scheme data)
	• vaccination status and coverage
	• health workforce data
	• prescriptions and medications
	• notifiable disease treatment records
	• public health management variables (e.g., health event such as COVID, infectious period, quarantine measures)
	
	Examples:
	• Australian Immunisation Register. Data items include vaccination coverage, immunisation provider and location, episode reason, vaccine product code, vaccination status. *[linked to PLIDA]*
	  ○ Data custodian: Department of Health and Aged Care
	• Medicare Benefits Schedule: Data items include Medicare claims information, benefit paid, broad type of service, number of services claimed, referring provider ID, patient demographic information, and bulk billing information. *[linked to PLIDA]*
	  ○ Data custodian: Department of Health and Aged Care
	
	Primary care, community, and other health data provided to the ABS via government and not-for-profit organisations is in scope of the PIA, and would be considered administrative health data.

2. Patient health and disease data	Includes information in clinical information systems, registers and databases that record data for chronic conditions, illnesses, and diseases. This includes cancer registry data and infectious and communicable disease data. This also includes outcomes of medical tests, biomedical markers and genetic testing for diagnosis of diseases or conditions. Data items may include: • cancer incidence • type of cancer • diagnostic information • disease type • mortality and survival • notification to heath authority • patient demographic information

Broad types of health data	Description and data items

	Examples:
	• Victoria Cancer Registry. Data items include patient information, cancer type and incidence, diagnosis information including notifying hospital and mortality. *[dataset being considered as one of the case studies]*
	  ○ Data custodian: Victorian Department of Health
	
	• National Notifiable Diseases Surveillance System (NNDSS), supplied by State and Territory health authorities to Australian Government Department of Health and Aged Care. Data items include disease code, date of onset, date of notification to health authority and patient information *[not linked to PLIDA, provided as example]*
	  ○ Data custodian: Department of Health and Aged Care

3. Health survey data	Includes information collected by clinical questionnaires, surveys, and studies of population health undertaken by government and/or not-for-profit organisations.
	
	Data items may include:
	• Health status
	• Medications and prescriptions
	• Injuries
	• Long term health conditions
	• Income and private health insurance
	
	Examples:
	• National Health Survey. Data items include chronic conditions, mental and behavioural conditions, general health status and health risk factors*. [linked to PLIDA]*
	  ○ Data custodian: Australian Bureau of Statistics
	
	• South Australian Population Health Survey. Data items include information on participants overall health status, health service utilisation, chronic conditions, disability and carers, risk factors, food security, mental health, wellbeing and disadvantage and inequity. *[not linked to PLIDA, provided as example]*
	  ○ Data custodian: SA Health
	
	• Victorian Population Health Survey. Data items include mental health and wellbeing status, chronic diseases, dental health, lifestyle risk factors and social determinants of health. *[not linked to PLIDA, provided as example]*
	  ○ Data custodian: Victorian Agency for Health Information (VAHI)

Broad types of health data	Description and data items
4. Ambulance and patient transport data	Includes information collected as part of responding to and treating out-of-hospital medical emergencies, patient transport services to assist patients and their families to attend appointments and transfers between heath facilities. Data items may include: • patient demographics • emergency medical care and interventions • pain management • facility transferred to (e.g., travel time, response locations and treatment time) • event type Example: • NSW Ambulance Patient Health Care Record. Data items include clinical information, patient care episodes, facility transfers, arrival time and location. *[linked to PLIDA]* ○ Data custodian: NSW Ministry of Health
5. Emergency department and outpatient data	Includes information about patients registered for care in emergency departments and responding to and treating out-of-hospital medical emergencies. Data items may include: • clinical records for non-admitted and emergency patients • screening or monitoring information • crisis and general counselling information • intervention type • medications and prescription information • Laboratory and diagnostic test and procedures (including pathology, x-ray or other medical imaging examinations) • reason for attendance • pain management Example: • QLD Emergency Department Data Collection. Data items include patient waiting times, urgency of care, presentation date, and complexity of presentations, patients who go on to be admitted to hospital or who are discharged and demand for ED services. *[linked to PLIDA]* ○ Data custodian: Queensland Department of Health
6. Hospital data	Includes information about care provided to admitted and non-admitted patients, including acute and non-acute care and mental health services. Data items may include: • patient demographics and physical characteristics • hospital episodes (e.g., length of stay, waiting times for care, admission date) • clinical information (e.g., patient diagnosis, interventions, and procedures) • hospital charges and health insurance information • admission, waiting times and length of stay. • perinatal care (mothers and babies)

Broad types of health data	Description and data items
	• obstetric records related to artificial insemination and in-vitro fertilisation (excluding donor records) Example: • Adult Patient Database. Data items include patient episodes, hospital type, ICU outcome, illness severity, diagnosis, and patient information. *[linked to PLIDA]* ○ Data custodian: Australia and New Zealand Intensive Care Society
7. Primary care, community, and other health data	Includes information collected from general practitioners, pharmaceutical services, dentistry, allied health services, maternal and child health, alcohol and drug treatment and other health related services. Data items may include: • Treatment illness and injury and service type • Attendance at general practitioners and specialist services • Diagnostic imaging information (including blood test and x ray results) • Mental health records • Dental records • Drug, gambling, smoking and alcohol data • Medication and prescriptions Example: • Public Dental Waiting times database. Data items include dental treatment, clinical assessment, number of appointments, area and waiting times for dental assessment. [*not linked to PLIDA, provided as example]* ○ Data custodian: Australian Institute of Health and Welfare (AIHW)
8. Out of scope	Specific Health Data out of scope: • My Health Record data • Private health insurance data • Clinical trial data • Genetic data (except outcomes of medical tests, biomedical markers and genetic testing for diseases of diseases or conditions) • Primary care, community and other health data collected directly from the health services provider

Attachment 3 Stakeholders consulted for the 2023/24 Expanded Health Data Linkage to the Personal Level Integrated Data Asset PIA

The list below sets out the organisations and other bodies that attended the consultation sessions for the Expanded Health Data Linkage PIA:[17]

1. Australian Bureau of Statistics (ABS)
2. Australian Capital Territory Department of Health
3. Australian Institute of Health and Welfare (AIHW)
4. Barang Regional Alliance - Empowered Communities (EC)
5. Cancer Alliance Queensland
6. Cancer Council Victoria
7. Cancer Institute New South Wales (CINSW)
8. Centre for Big Data Research in Health University of NSW
9. Consumers Health Forum of Australia
10. Data Linkage Queensland
11. Department of Employment & Workplace Relations (DEWR)
12. Department of Health & Aged Care (DOHAC)
13. Department of Home Affairs (DoHA)
14. Department of Social Services (DSS)
15. Disability Advocacy Network Australia (DANA)
16. Empowered Communities (Independent) – Independent Social Alchemist
17. Families Australia
18. Grattan Institute
19. Inclusion Australia
20. Indigenous Data Network
21. Kimberley Aboriginal Medical Services (KAMS)
22. Mental Health Australia
23. Monash University
24. Northern Territory Cancer Registry
25. Northern Territory Department of Health
26. Northern Territory Office of the Information Commissioner
27. Office of the Australian Information Commissioner (OAIC)
28. Population Health Research Network (University of Western Australia)
29. QLD Health in Thursday Islands and QLD Peninsular communities
30. Safer Care Victoria
31. South Australian Department of Health
32. Tasmanian Data Linkage Unit
33. Tasmanian Department of Health
34. The Peter Doherty Institute for Infection and Immunity
35. University of Sydney
36. Victorian Agency for Health Information
37. Victorian Cancer Registry
38. Victorian Department of Health
39. Western Australian Department of Health

[17] One or more individuals from the listed entities attended the consultation sessions, but views provided by those individuals may not necessarily have been representative of the views of their entity. Several other organisations and other bodies were also invited to participate in the consultation sessions but were not able to attend.