

# Privacy Impact Assessment

## Australian Immunisation Register (AIR) Data Integration Project

### August 2021

#### 1. Introduction

The ABS has conducted a Privacy Impact Assessment (PIA) for the integration of the Australian Immunisation Register (AIR) with the Multi-Agency Data Integration Project (MADIP). Integration of AIR data with MADIP is being used to inform the Government's COVID-19 Vaccine and Treatment Strategy. The integrated data will enable Government researchers to analyse vaccine take-up by a range of socio-demographic cohorts and inform targeted policy interventions to achieve a high-level of vaccination.

The AIR dataset includes information relating to immunisation status for COVID-19 and other vaccines, and basic core demographic information (e.g. age and sex). The integrated MADIP-AIR dataset combines AIR data with existing MADIP data including: National Disability Insurance Scheme (NDIS) provider and participant data; aged care provider data; personal tax information related to occupation and industry; migration status; health data; Single Touch Payroll data and Centrelink data.

The AIR data was first linked to MADIP as a 'once-off' linkage for the Department of Health in early 2021 and then on a fortnightly basis up until mid-July 2021. From mid-July, AIR data has been supplied to the ABS weekly, with Government researchers provided access to data just over a week later. The AIR-MADIP linkage has now become an enduring MADIP asset, with other researchers able to request access to the de-identified microdata for statistical and research purposes.

Before acquiring the AIR data in early 2021, the ABS undertook a privacy threshold assessment (PTA) for the AIR linkage with MADIP. The assessment identified that the proposed linkage was not a high privacy risk project. It did not introduce any new types of data to those already integrated through MADIP or involve any changed data handling practices which had not been assessed as part of previous MADIP PIAs. The PTA did recommend that if researchers were to access AIR data relating to time periods less than two weeks old, a PIA update should be undertaken.

This PIA considers the privacy impacts of the AIR-MADIP integration, specifically focusing on the increased recency of data now supplied, and outlines strategies to mitigate those impacts.

#### 2. Information and information flows

##### Type of information

The Privacy Act defines personal information as, "...information or an opinion about an identified individual, or an individual who is reasonably identifiable...".

The ABS recognises the sensitivity around personal health information. The integrated AIR-MADIP data includes detailed information about people living in Australia, such as whether they have received a COVID-19 vaccination, whether they work in aged care or NDIS support services, and whether they live in aged care or disability support homes. The integrated data includes several sources of data that, if accessible with personal identifiers, would be considered “sensitive information” under the Privacy Act 1988.

Standard MADIP data integration infrastructure and arrangements are used to integrate the AIR data and minimise the handling and use of personal information. All data accessed by researchers is de-identified in accordance with ABS microdata access standard processes, meaning an individual will not be reasonably identifiable from the data, and the data accessed by researchers will therefore no longer be personal information. Researchers are able to apply for access to sensitive data for specific research proposals. Requests will be considered using standard MADIP data access arrangements, including a requirement to justify access to all sensitive data. Further, access is through a secure, closed ABS system to remove the potential for data to be re-identified whilst research is underway as well as in the resulting research outputs. These standard processes and procedures have been assessed previously as part of the MADIP PIAs conducted in 2018 and 2019, available on the [ABS Privacy Impact Assessment webpage](#).

### **AIR Information flow**

AIR data is extracted from the Department of Health’s Enterprise Data Warehouse. The extract is encrypted and transferred to the ABS using the Department’s Health Data Portal. This data transfer system is highly secure and subject to regular internal and external security assessments. Once the data has been transferred, it is securely stored within the ABS environment for data integration purposes.

### **MADIP Data Asset, infrastructure and arrangements**

The AIR-MADIP project uses existing MADIP governance and infrastructure to ensure the secure integration of the AIR data with other data assets. Standard MADIP information flows and data handling procedures are discussed in detail in the *2019 MADIP PIA Update*.

MADIP is a partnership among Australian Government agencies and operates under a framework of legislation, governance, and information management protocols that ensures data are shared and used for public benefit, privacy is protected, and data remains secure.

Data custodians, or entities authorised by data custodians, share data with the ABS for MADIP. The ABS is the accredited Integrating Authority for MADIP and is responsible for receiving, storing, and linking data, as well as assembling extracts of integrated data for analysis and providing access to these extracts to authorised researchers. Access to de-identified integrated AIR analytical microdata data can be requested by approved researchers for approved projects via the ABS DataLab. In line with the MADIP operating model, data custodian approval will be sought for any access requests.

MADIP data is covered by the legislative protections of the Census and Statistics Act 1905. Data can only be used for statistical or research purposes, and outputs must be consistent with the requirements of the Census and Statistics Act.

### 3. Privacy risk and analysis

The linking of a wide range of types of data in MADIP, including fortnightly updates for a range of data sources, was considered in other data integration PIAs including the *2019 MADIP PIA Update* and the *2020 Jobs Related Data Integration PIA*<sup>1</sup>. The Jobs PIA notes that, in general, the older the data is, the harder it is to identify people correctly from it. This is because people's information may change over time and older data may not be easily or accurately recalled. It concluded that MADIP data integration arrangements and processes provide sufficient protection for integrated data older than two weeks to be used by the authorised researchers in the ABS DataLab.

This PIA specifically examines the potential risk of re-identification and unauthorised disclosure of personal information through accessing AIR data that is less than two weeks old. The assessment is being made against Australian Privacy Principle 11 – Security of personal information – and the risk of unauthorised disclosure of, or access to, personal information.

Access to data that is updated frequently and in near-real time may present a slightly greater risk of re-identification of personal information, especially where researchers have knowledge of recent immunisation events. The ABS identified several factors that may pose an increased risk of identification of an individual including:

- The early stages of a vaccination roll out with relative low numbers of vaccinations
- New vaccines with low numbers of vaccinations in early stages of use
- Low participation rates and numbers from specific cohorts (e.g. children) especially when the scope of eligibility changes and
- Where information contains unique or uncommon characteristics (such as a high-profile person immunised on a particular date) where vaccination numbers are low.

The ABS applied its standard processes to minimise the privacy risk, including data specific controls to strengthen privacy protections and treat the data risks. These standard treatments will effectively minimise disclosure risk. Any residual risk of researcher re-identification of personal information will be managed using the Five Safes Framework as discussed in detail in the *2019 MADIP PIA Update*.

The source data supplied by the Department of Health will be stored in accordance with the separation principle, ensuring that personal identifiers (linkage data such as names, addresses and

---

<sup>1</sup> Both PIAs were conducted with independent privacy advice, review and assurance including that the processes were consistent with the Office of the Australian Information Commissioner (OAIC) Guide to undertaking privacy impact assessments. A range of stakeholders, including the OAIC, were consulted as part of the PIA processes.

scrambled government identifiers) cannot be accessed with analytical data. Key features to note include:

- Separate transfer – linkage and analytical data is transferred to the ABS in separate files.
- Functional separation – linkage and analytical data cannot be accessed together by anyone at any time.
- Regular auditing, access controls and permissions – ensuring that functional separation is properly applied.

#### 4. Summary

The *2019 MADIP PIA Update* assessed MADIP standard processes and arrangements against all the APPs. This PIA has considered any additional privacy risks associated with the linkage of AIR data with MADIP and researcher access to near real-time immunisation data (see section 3).

Based on this assessment, the ABS has determined that the standard MADIP procedures for handling data for the purpose of integration provide sufficient protections for the frequent and timely AIR-MADIP data integration. In particular, MADIP disclosure risk mitigation strategies are appropriate and sufficient to address the disclosure risk from access to recent (less than two weeks old) immunisation data. As a result, this PIA does not make any further recommendations.

The ABS enforces a robust framework of protections that work together to protect the privacy and security of MADIP data and meets all legislative requirements. The AIR-MADIP data integration project will have the full protections of MADIP infrastructure and arrangements to ensure the data is kept safe.

The ABS is planning to undertake a MADIP Privacy Impact Assessment Update in late 2021. This update will consider the general case of integration of near-real time data in MADIP. The PIA process will involve broad stakeholder consultation to seek feedback around risk arising from increased frequency of data for integration.