Government of Western Australia
Department of Health

Guideline: Privacy Preserving Record Linkage at the Department of Health

Classification: Official

Table of Contents

1. Background 2
2. Purpose and scope 2
3. Development of a PPRL service 2
4. Implementation of PPRL 4
5. Roles and responsibilities 7

Stage 1: Pre-application 7
Stage 2: Preliminary review 7
Stage 3: Data Custodian Feasibility Assessment 8
Stage 4: Ethical approval 10
Stage 5: Data Custodian Formal Review 11
Stage 6: Governance approvals 12
Stage 7: Cohort pool selection 13
Stage 8: Data linkage 14
Stage 9: Data extraction 14
Stage 10: Data delivery 15
Stage 11: Invoicing 15
Stage 12: Project closure 16

6. References 17
7. Acronyms and Definitions 19
8. Appendices 21
9. Document Control 23

1. Background

Traditional probabilistic linkage methods which utilise full clear text personal identifiers is the preferred linkage method of the Department of Health (the department), as it delivers higher accuracy, stronger confidence in match results, and better handling of real-world data imperfections. The department has established and maintains a strong commitment to protecting information from misuse or inappropriate disclosure.

There are instances where legislative barriers and other limitations prohibit the sharing of identifiable data for data linkage. Privacy preserving record linkage (PPRL) method can be a solution for data providers who have concerns about privacy and security or are restricted from sharing clear text demographic information for linkage.

PPRL can enable data linkage without the disclosure of personally identifying information. Fields such as name, address and date of birth are irreversibly 'hashed' using specialised software to derive new strings that still enable record linkage but do not identify an individual.

PPRL supports the sharing and integration of data where legislative barriers and other limitations prohibit the sharing of identifiable data. This supports government service delivery, research, and other initiatives while adhering to the Australian Privacy Principles.

2. Purpose and scope

The purpose of this document is to provide guidance to parties by:

- Outlining the roles and responsibilities of stakeholders involved in facilitating access to data via the PPRL method, and
- Illustrating processes within the department to securely link information that is subject to legislative and policy restrictions, or security concerns.

The scope of this document is limited to the PPRL service provided by the department in collaboration with third-party linkage agencies and does not apply to other data services provided by the department.

This document is intended to support data applicants with their understanding of the PPRL process and promote transparency of the department's PPRL service.

3. Development of a PPRL service

The department collaborated with the Centre of Data Linkage (CDL) at Curtin University to conduct three Proof of Concept (PoC) PPRL projects.

These projects assessed the feasibility of using a third-party agency to undertake privacy-preserving linkage of data received from providers who are unable to provide personally identifiable data to the department for traditional clear-text linkage.

In PPRL data linkage, the identifying data from all relevant collections is hashed by the appropriate data provider using specialised software (e.g. LinXmart), which uses field-based filters known as Bloom filters. Bloom filters create multiple hashes of each field,

which allows for links to still be made in the presence of textual variants and misspellings. Hashed data is securely delivered to the third-party linkage agency, who links encoded data from all data providers.

The proof-of-concept projects established that, provided the identifying data in each data collection was of reasonable quality, there was very little difference in the linkages that were created through PPRL and clear-text linkage, showing that PPRL is a viable option that can produce good quality linked data when clear text linkage is not possible. However, further work is needed to better understand how PPRL performs in contexts where demographic information may vary for cultural or community reasons. Examples include temporary name changes, the use of generic or placeholder dates of birth, and scenarios in which multiple non-familial individuals share an address.

Following the close-out of this project, the department has established a PPRL service model, which can follow one of two processes, depending on the project: 'Hybrid' and 'Standalone' PPRL.

The primary difference between the two processes is whether the linkage key generated by the WA Data Linkage System (WADLS) is provided to the third-party linkage agency alongside the hashed data, enabling the linkage agency to leverage clear-text probabilistic linkage performed by the department, whilst still maintaining privacy and data security.

All data updates and amendments requiring extraction of additional records (e.g. more recent data) provided by the department will require a full refresh for both Hybrid and Standalone PPRL.

More detail of each process is provided below.

A. Hybrid PPRL

Linkage keys for datasets held by the WADLS can be provided by the department to the third-party agency to leverage the benefits of both clear text probabilistic linkage and linkage supported by the routine integration of high-value datasets within an enduring, broad scope linkage system. These benefits, in turn, improve linkage quality across health-provided datasets.

Using the WADLS linkage key also allows access to linked health service data that predates the cohort pool and has not been linked using PPRL directly. This means that, depending on the project, PPRL linkage may not be required for the entire requested date range, potentially streamlining processes and reducing the amount of hashed data provided to the third-party linkage agency.

However, for Hybrid PPRL to be feasible:

- The third-party linkage agency must accept the WADLS linkage key as valid; that is, records provided to the third-party agency with the same WADLS linkage key must also be returned to the department with the same linkage key. This is important as discrepancies between the WADLS linkage key and the PPRL linkage key for health records may lead to errors in the linked content data.
- Projects that use the Hybrid PPRL approach must acknowledge the use of clear-

text probabilistic linkage performed by WADLS as part of the project's overall linkage process, highlighting its role in producing higher quality linked data.

In the Hybrid PPRL method, the department is responsible for links created by clear-text probabilistic linkage and the third-party agency is responsible for links created by PPRL. As such, both the department and the third-party agency may respond to applicant queries after the integrated dataset is delivered, depending on the nature of the query.

The quality of links in the WADLS is supported by clerical reviews, manual checks and quality assurance procedures implemented by the Data Linkage team at the department.

B. Standalone PPRL

In standalone PPRL, the WADLS linkage key is not shared with the third-party linkage agency and PPRL is used for all linkage performed as part of the project.

Projects using standalone PPRL can be considered as relying exclusively on PPRL linkage, with no clear-text probabilistic linkage involved.

To ensure all relevant health records linked by PPRL can be included in the integrated dataset, the hashed data delivered to the third-party agency must encompass all records that may be part of the final integrated dataset.

The third-party linkage agency responsible for facilitating PPRL is fully accountable for all aspects of the linkage process, including quality assurance and responding to applicant queries following the delivery of the integrated dataset.

4. Implementation of PPRL

To implement PPRL as a service, the following measures have been introduced by the department.

Project Management Tool

The department utilises a project management tool to oversee the different stages of the data delivery pipeline. A specific master template has been created for the PPRL data delivery pipeline, incorporating the additional steps required for this linkage method.

Application Form

The department has developed an application form tailored for PPRL projects.

External Data Checklist

As part of the process, all external data providers must complete a checklist (Appendix 2) to ensure that the provided data aligns with the project's specifications before being encrypted and linked using PPRL.

The third-party linkage agency facilitating PPRL will distribute this form to the relevant data providers as part of the ordinary workflow.

Standard Operating Procedures

To streamline PPRL as a regular service within the department, two Standard Operating Procedure (SOP) documents have been developed.

The first SOP focuses on the extraction of data for PPRL projects, outlining unique variations and dependencies specific to the PPRL process and presenting different approaches to data delivery for PPRL projects.

The second SOP documents the process of data hashing and generation of hashing keys using LinXmart, a client tool provided by the third-party linkage agencies that will be facilitating PPRL.

Project Amendments for Data Updates

Under the first revision of this guideline no provision was included for updates to data provided with PPRL.

Project amendments involving updates or refreshes to data provided by the department will require a full refresh of these datasets. This requirement applies to both Hybrid and Standalone PPRL projects.

- For Hybrid PPRL projects, the dynamic nature of WADLS keys necessitates a complete refresh to maintain linkage integrity.
- For Standalone PPRL projects, current technical limitations prevent the reliable management of incremental changes (“deltas”), making a full refresh the only viable option.

However, Data Outputs will design and build a process to provide delta data for standalone PPRL projects in the future.

Information Sharing Agreement

To facilitate the sharing of information while protecting information from misuse or inappropriate disclosure, an Information Sharing Agreement (ISA) may be developed between the department and a third-party linkage provider for the purposes of PPRL service delivery.

The department has developed an ISA to sit alongside the previously signed Memorandum of Understanding (MOU) with Curtin University on 12 January 2022, governing data sharing between the department and CDL for PPRL.

The ISA aims to:

- optimise the value, quality, and security of information;
- ensure compliance with applicable laws and regulations;
- minimise misuse and inappropriate disclosure of information;
- promote appropriate governance mechanisms, and foster information integrity.

Schedule 2, which includes the Information Request Agreement form, will be completed on a project-by-project basis by the department’s ISPD Client Services branch, regardless of the PPRL third-party provider.

Utilisation of Separation Principle

All stakeholders will adhere to best practice protocols, including the separation principle, to safeguard the privacy of individuals whose personal information is used or disclosed.

The third-party PPRL service provider implements this principle by keeping encoded

identifiers separate from content data (i.e., event-level service information connected to an individual's record) and maintaining equivalent security and governance processes for encoded and un-encoded data.

It is a requirement for third-party agencies providing PPRL services to operate an ISO-certified information security management system in a physically secure environment, maintaining strict access controls when handling data, including receiving data extracts or hashed data.

Further details on processes are included within the process map in Appendix 1.

5. Roles and responsibilities

Collaboration between the applicant, third-party linkage agencies, and the department is necessary to advance a project via PPRL.

If a project necessitates the use of PPRL for linkage, it will go through the standard application process at the department as laid out within the Research Data Services Roles and Responsibilities guide.

Research projects require involvement of several stakeholders, including the applicant, ISPD Client Services request coordinators, data custodians, the Research Governance Office (RGO), data engineers, and ethical bodies.

Stage 1: Pre-application		
Step	**Party**	**Task**
1.1	Applicant	Contact ISPD Client Services regarding data and application requirements.
1.2	ISPD Client Services	Provide advice to applicant (as required). Seek feedback from Data Custodians on data queries and the Data Linkage and Data Engineering teams regarding the technical viability (if needed).
1.3	Applicant	Contact RGO to determine the project's viability and application requirements in terms of governance and ethics.
1.3.1	RGO	Provide advice to applicant on ethical and governance requirements.
1.4	Applicant	Submit draft data application to ISPD Client Services via email for a preliminary review.
1.4.1	**MILESTONE**	**DRAFT APPLICATION SUBMITTED**

Stage 2: Preliminary review		
Step	**Party**	**Task**
2.1	ISPD Client Services	Administrative processing.
2.1.1	ISPD Client Services	Contact third-party linkage agencies that will be facilitating PPRL to advise that a research application requiring PPRL has been received and is currently being reviewed.
2.2	Applicant	Apply for a login to the Research Governance Service (RGS) system to commence ethics and research governance process. Complete the New User Sign Up form and submit to the RGS Administrator for approval (if needed). If applicant is not a WA Health employee, they will also need to provide a WA Health employee as a referee/Local Contact in the form to verify the user login

		request. This can be the relevant ISPD Client Services Request Coordinator.
2.2.1	RGS Administrator	If the applicant is not a WA Health employee, the RGS Administrator will contact the WA Health employee who has been nominated as the local contact to verify the applicant's identity and confirm the requirement for access to RGS.
2.2.1.1	ISPD Client Services	Verify the user login request (if applicable).
2.3	RGS Administrator	Approve user (applicant) request for access.
2.4	Applicant	Create a project workspace in RGS.
2.5	RGS Administrator	Approve the project workspace in RGS.
2.6	ISPD Client Services	Conduct a preliminary review of the draft data application to ensure completeness, and assign a priority to the project in accordance with the Data Linkage Prioritisation and Costing Framework.
2.7	ISPD Client Services	Collate feedback and send to applicant.
2.8	Applicant	Review feedback, action necessary changes and resubmit draft data application (as required).
2.9	ISPD Client Services	Additional preliminary review of draft data application (as required). If feedback is adequately addressed, project is added to the agenda for the next available draft data application review meeting. This is an iterative process which may require multiple reviews.
2.10	ISPD Client Services	Administrative processing.
2.10.1	ISPD Client Services	Send the draft data application to Data Custodians for review ahead of the Draft Data Application Review meeting.
2.10.2	ISPD Client Services	Send the draft data application to Data Linkage, Data Engineering for technical feasibility review ahead of the Draft Data Application Review meeting. The application may also be sent to third-party linkage agencies facilitating PPRL for feedback, if necessary.
2.11	ISPD Client Services	Create data flow diagram specific to the project.

Stage 3: Data Custodian Feasibility Assessment		
Step	**Party**	**Task**
3.1	**MILESTONE**	**FEASIBILITY ASSESSMENT**
3.1.1	Data Custodians	Review draft application for feasibility and against data use and disclosure requirements, including relevant legislation, data sharing agreement and policy frameworks which apply to the data asset in question.

3.1.2	Data Custodians	Review draft application to determine if the data requested is fit for purpose and can satisfy the objectives of the project.
3.1.3	Data Linkage; Data Engineering Outputs; Third-party linkage agencies facilitating PPRL	Conduct Technical Feasibility Assessment and provide feedback, recommendations and/or comments for the applicant in preparation for the Draft Application Review meeting. Confirm whether Hybrid PPRL or standalone process is required.
3.1.4	RGO	Conduct governance feasibility review and provide feedback, recommendations and/or comments for the applicant in preparation for the Draft Data Application Review meeting.
3.1.5	WADLS Custodian	Review linked data requests against data use and disclosure requirements, including relevant legislation, data sharing agreement and policy frameworks, and provide feedback, recommendations and/or comments with respect to the use of linkage keys and linkage products in alignment with their permitted and suitable use (where relevant).
3.2	**MILESTONE**	**DRAFT APPLICATION REVIEW MEETING**
3.3	ISPD Client Services	Provide the applicant with consolidated feedback from Data Custodians, Data Linkage, Data Engineering Outputs, and third-party linkage agencies that will be facilitating PPRL. Indicate to the applicant if Hybrid PPRL is intended to be used by third party linkage agencies and that acknowledgement is required for PPRL hybrid solutions (see Section 3A of Guide).
3.3.1	Applicant	Review feedback provided by ISPD Client Services and resubmit revised draft data application. This is an iterative process which may require multiple reviews.
3.3.2	ISPD Client Services	Process and check revised draft data application, before providing to relevant Data Custodians for additional review (as required).
3.3.3	ISPD Client Services	Coordinates technical feasibility review between third-party linkage agencies that will be facilitating PPRL, Data Linkage and Data Engineering - Data Outputs. Provide the PPRL data checklist for third-party linkage agencies that will be facilitating PPRL to circulate to external data providers.

3.4	MILESTONE	DATA CUSTODIAN IN-PRINCIPLE SUPPORT
3.4.1	Data Custodians	Once satisfied that feedback has been addressed, advise ISPD Client Services of in-principle support.
3.5	ISPD Client Services	Calculate cost estimate for department services only in accordance with the Data Linkage Prioritisation and Costing Framework..
3.5.1	ISPD Client Services	Inform third-party linkage agencies facilitating PPRL that the cost estimate is being calculated and will be provided to applicant directly.
3.5.2	Third-party linkage agencies facilitating PPRL	Third-party linkage agencies facilitating PPRL to calculate the cost estimate for PPRL services only and provide the cost estimate and details of how to accept it directly to the applicant.
3.6	ISPD Client Services	Prepare feasibility letter confirming in-principle support from Data Custodians to enable applicant to proceed to the WA Health Central HREC for review (ethics application).
3.6.1	ISPD Client Services	Provide cost estimate and feasibility letter to the applicant confirming in-principle support from Data Custodians for project to proceed ethics application. Include the latest data application form and data management plan in the email. Inform the applicant that third-party linkage agencies facilitating PPRL will be providing a separate cost estimate for the PPRL services.
3.6.2	Applicant	Upload final data application form, data management plan, variable lists, extraction forms and feasibility letter into RGS.

Stage 4: Ethical approval		
Step	**Party**	**Task**
4.1	Applicant	Contact RGO and inquire on all required ethical approvals as well as clarifications on whether a Data Transfer Agreement (DTA), an ‘Access request Form’, or ‘Site Specific Assessment Form’ are required for the project.
4.2	Applicant	Submit ethics application and the final approved data application documents to the WA Health Central HREC for review and approval via RGS.
4.3	Ethics Executive Officer	Review ethics application to ensure completeness and validity, requesting further information from applicant if required.
4.3.1	Ethics Executive Officer	If feedback is adequately addressed, project is added to the agenda for the next available WA Health Central HREC meeting.

4.4	WA Health Central HREC	Discuss and review research ethics application.
4.4.1	Applicant	If a significant amendment to the ethics application is required following ethical review, the applicant may have to revert to Stage 3 and resubmit draft data application.
4.4.2	Applicant	If application requires minor amendments, the applicant can amend ethics application and resubmit for ethical approval (Step 4.1).
4.5	Other HRECs	If applicable, the applicant must submit an additional ethics application to other HRECs for review and approval.
4.5.1	Applicant	If a significant amendment to the ethics application is required following ethical review, the applicant may have to revert to Stage 3 and resubmit data application.
4.5.2	Applicant	If application requires minor amendments, the applicant can amend ethics application and resubmit for ethical approval (Step 4.4).
4.6	**MILESTONE**	**ETHICS APPROVAL GRANTED**
4.6.1	Ethics Executive Officer	Provide letter to the applicant confirming ethical approval via RGS within five (5) working days of the WA Health Central HREC meeting.
4.6.2	Other HRECs	If applicable, provide letter to the applicant confirming ethical approval.
4.6.3	Applicant	Upload other HREC approval letter/s (if applicable) via RGS.

Stage 5: Data Custodian Formal Review		
Step	**Party**	**Task**
5.1	WA Health Central HREC	Notify applicant of ethical approval via email with the following attachments: • Data application approved by HREC • HREC approval letter
5.1.1	Applicant	Applicant to notify ISPD Client Services of ethical approval.
5.2	ISPD Client Services	Cross-reference the HREC-approved data application against the IPS application to identify any significant changes
5.3	ISPD Client Services	Request formal Data Custodian approval from all Data Custodians/Stewards.
5.3.1	ISPD Client Services	Advise the relevant Data Custodians of any significant changes to the application between in-principle support and HREC approval.

5.3.2	ISPD Client Services	Advise third-party linkage agencies facilitating PPRL of any technical changes or variations to the proposal made during the ethical review process.
5.4	Data Custodians	After ethical approval(s) are finalised and all relevant criteria satisfied, review the WA Health Central HREC-approved data application forms and provide formal Data Custodian approval.
5.5	**MILESTONE**	**DATA CUSTODIAN FORMAL APPROVAL GRANTED**
5.5.1	ISPD Client Services	Notify applicant that formal approval has been granted, including advice to engage with RGO regarding research governance approval application requirements, and to add ISPD Client Services Request Coordinator as 'Data Custodian' on the Access Request Form or Site-Specific Assessment Form and Budget Form as applicable.
5.5.2	ISPD Client Services	Contact third-party linkage agencies facilitating PPRL to ensure Information Sharing Agreement (ISA) Schedule 2 is completed before hashed data is provided for linkage. Schedule 2 of the ISA will be completed on a project-by-project basis.
5.5.3	Third-party linkage agencies facilitating PPRL	Start contractual agreement with Applicant to deliver PPRL service. Contractual agreement must note whether PPRL or PPRL hybrid method will be used.

Stage 6: Governance approvals		
Step	**Party**	**Task**
6.1	Applicant	Ensure all ethic approval letter/s (if applicable) have been uploaded into RGS.
6.2	Applicant	Liaise with RGO to complete the draft DTA if the applicant is external to the WA Health system.[1] DTA template can be obtained from RGO.
6.3	Applicant	After being advised of Data Custodian formal approval, submit the DTA (if applicable) and relevant site authorisation form via RGS, and add the relevant Request Coordinator as the 'Data Custodian' (if applicable).
6.3.1	ISPD Client Services	Approve the applicant's Access Request Form submission (if applicable) to trigger RGO review, and upload Data Custodian Approval for Research Project document via RGS.

[1] Per section 19 of the Health Services Act 2016, the WA Health System is comprised of:
(a) the Department of Health; and
(b) health service providers; and
(c) to the extent that contracted health entities provide health services to the State, the contracted health entities.

6.4	Applicant	Project Principal Investigator to sign the Access Request, Site Authorisation, and Budget form(s) and complete the submission to the department's RGO via RGS.
6.5	RGO	Review DTA (if applicable) and request further information from the applicant (as required).
6.5.1	Legal and Legislative Services; Insurance Commission of Western Australia	Significant changes to a DTA (if required) are referred to Legal and Legislative Services (LLS) and/or the Insurance Commission of Western Australia (ICWA) for review and/or advice as appropriate.
6.6	RGO	If the project requires a DTA, prepare a Briefing Note for approval by the delegated authority.
6.7	Director, Information and Performance Governance (IPG)	Reviews and approves the DTA.
6.8	RGO; Applicant	Complete site authorisation process within 60 days of receiving request from applicant, notwithstanding time taken for feedback and requests for further information.
6.9	RGO	Write a recommendation to the Deputy Director General (DDG) (or delegate) that they consider providing site authorisation for the project and for data to be released.
6.10	DDG (or delegate)	Use RGS to indicate if site authorisation is (or is not) granted for the project.
6.11	**MILESTONE**	**GOVERNANCE APPROVED**
6.11.1	RGO	Provide finalised DTA (if applicable) and letter to the applicant confirming site authorisation via RGS
6.11.2	RGO	Notify ISPD Client Services of research governance approval and provide a copy of site authorisation via email

Stage 7: Cohort pool selection		
Step	**Party**	**Task**
7.1	ISPD Client Services	Prepare for project kick-off and schedule data linkage and extraction work in project queue.
7.2	Data Linkage; ISPD Client Services; Data Engineering Outputs	Evaluate request and project requirements with support from ISPD Client Services and Data Engineering Outputs as required.
7.3	Data Engineering Outputs; Data Linkage; ISPD Client Services	Cohort pool selection planning meeting.

7.4	Data Engineering Outputs	Perform cohort pool selection and provide files to Data Linkage Perform service data extraction for entire cohort pool and save as snapshot for later use in stage 9.

Stage 8: Data linkage		
Step	**Party**	**Task**
8.1	Data Linkage	Receive and validate cohort pool files from Data Engineering Outputs
8.2	Data Linkage; ISPD Client Services; Data Engineering	Linkage project planning meeting.
8.3	Data Linkage	Generate hashing key and provide to external data provider(s).
8.4	Data Linkage	Perform extraction and hashing of demographic data for cohort pool (from step 7.3). This task is performed within the WADLS.
8.5	Third-party linkage agencies facilitating PPRL	Confirm with ISPD Client Services that the applicant has signed the contract for PPRL services (if required).
8.6	Data Linkage	Send hashed Data to third-party linkage agencies that will be facilitating PPRL (pending confirmation of Step 8.5 above). If following a hybrid PPRL process, include encrypted WADLS linkage keys within the hashed data provided to the third-party linkage agency facilitating PPRL. Notify ISPD Client Services that hashed data has been sent.
8.7	Third-party linkage agencies facilitating PPRL	Perform PPRL.
8.8	ISPD Client Services	Provide a project status update to the applicant via email.
8.9	Data Linkage	File all linkage-related project email correspondence as per standard recordkeeping process.

Stage 9: Data extraction		
Step	**Party**	**Task**
9.1	Third-party linkage agencies facilitating PPRL	Provide README document and linkage mapping file for relevant records to be extracted to Data Engineering.
9.2	Third-party linkage agencies facilitating	Email to Data Outputs and ISPD Client Services to confirm that the data has been sent.

	PPRL	
9.3	Data Engineering Outputs; Data Linkage; ISPD Client Services	Project planning meeting for clarification of data request.
9.4	Data Engineering Outputs	Data extraction of requested datasets (using snapshot created in step 7.4). This data extraction function is limited to WA Health datasets.
9.5	Data Engineering	Quality assurance review. For hybrid projects, this includes ensuring the WADLS linkage keys are not in the final extracts.
9.6	ISPD Client Services	Provide a project status update to the applicant via email.
9.7	**MILESTONE**	**LINKAGE AND EXTRACTION COMPLETE**
9.8	ISPD Client Services	Administrative processing.

Stage 10: Data delivery		
Step	**Party**	**Task**
10.1	Data Engineering Outputs	Provide data extractions to ISPD Client Services.
10.1.1	Data Engineering Outputs	Include mapping file provided by third-party linkage agencies facilitating PPRL with the final data extract.
10.2	ISPD Client Services	Prepare data for delivery to applicant, including encryption of the data.
10.3	ISPD Client Services	When all required approvals are complete and the approved recipient has confirmed contact details, release data to the recipient via secure online file transfer or access platform.
10.4	**MILESTONE**	**DATA DELIVERED**
10.4.1	Applicant	Receive data via secure online file transfer or access data via secure access platform once all documents have been approved.

Stage 11: Invoicing		
Step	**Party**	**Task**
11.1	ISPD Client Services	Calculate the final cost to charge the applicant (if applicable). This task is restricted to Department of Health services only.
11.1.1	ISPD Client Services	Investigate any significant changes from the cost estimate.

11.2	ISPD Client Services	Send final cost to applicant with explanation for any significant variances from initial cost estimate via email (if applicable).
11.2.1	Third-party linkage agencies facilitating PPRL	Provide separate Final invoice to applicant as per agreement with third-party linkage agencies facilitating PPRL (see point 3.5.2).
11.3	Applicant	Review and confirm final cost. Provide Purchase Order number.
11.4	ISPD Client Services	Prepare S60 Debtor Advice Form for progression to Department's Accounts Receivable team to generate invoice.
11.5	Accounts Receivable	Issue invoice to applicant's designated contact.
11.6	Applicant	Pay invoice within designated timeframe.
11.7	Accounts Receivable	Notify Client Services of projects with outstanding payments.
11.8	ISPD Client Services	If payment hasn't been received, follow up with the applicant.

Stage 12: Project closure		
Step	**Party**	**Task**
12.1	RGO	Site-specific monitoring of the conduct of research projects, to ensure that projects RGO authorise are conducted with integrity and in compliance with relevant procedural requirements.
12.2	Applicant	Provide annual progress reports to the WA Health Central HREC via RGS (or as per HREC conditions of approval).
12.2.1	WA Health Central HREC	Monitor the applicant's research to ensure approved projects are conducted ethically in adherence with relevant conditions.
12.3	Applicant	Submit Site Final Report to RGO via RGS after the completion of the project analysis.
12.3.1	RGO	Validate, review, and authorise/approve the Site Final Report submitted by the applicant via RGS.
12.4	Applicant	Upon approval of Site Final report, submit Project Final Report via RGS.
12.4.1	WA Health Central HREC	Validate, review, and authorise/approve the Project Final Report submitted by the applicant via RGS.
12.5	ISPD Client Services	Send Client Feedback Survey to applicant to assess the quality of services provided by the department.
12.6	Applicant	Provide a copy of project draft outputs (study article, etc.) to ISPD Client Services via email for review prior to publication.

12.7	ISPD Client Services; Data Custodians	Coordinate review of project draft outputs and complete Draft Output Review Checklist. ISPD Client Services to liaise with Data Custodians (if applicable) to review project draft outputs.
12.7.1	ISPD Client Services	If the applicant has requested to publish small cell counts in their output(s), submit the output(s) to the department's HREC for review in the monthly meeting.
12.7.2	WA Health Central HREC	If applicable, review draft output(s) and provide outcomes to ISPD Client Services.
12.8	ISPD Client Services	Provide feedback to the applicant regarding publication of project outputs.
12.9	Applicant	Provide a copy of project outputs (study article, etc.) to ISPD Client Services after publication.
12.10	All (excl. applicant)	Ensure all project records are saved into the Electronic Document and Records Management System in adherence with the *State Records Act 2000.*
12.11	ISPD Client Services	Liaise with third-party linkage agencies to seek review of outputs or publications from the project.

6. References

Legislation and Regulations

- *Health Services Act 2016* (WA) and *Health Services (Information) Regulations 2017* (WA)
- *Freedom of Information Act 1992* (WA)
- *Privacy and Responsible Information Sharing Act 2024* (WA)
- *Privacy Act 1988* (Cth) and *Privacy Amendment (Enhancing Privacy Protection) Act 2012* (Cth)
- *State Records Act 2000* (WA)

Policies and Procedures

- Information Management Policy Framework
    - MP 0015/16 Information Access, Use and Disclosure Policy
        - Information Access, Use and Disclosure Standards
- Research Policy Framework
    - MP 0162/21 Research Governance Policy
        - Research Governance Procedure

Guidance Resources

- Research Data Services Roles and Responsibilities Guideline
- Information Access, Use and Disclosure Policy Resource Compendium
- WA Health Central HREC resources
- Australian Privacy Principles
- Practice Code for the Use of Personal Health Information
- IT Security Risks and Mitigating Controls for Research
- *Australian Code for the Responsible Conduct of Research* (2018)
- *National Statement on Ethical Conduct in Human Research* (2023)

7. Acronyms and Definitions

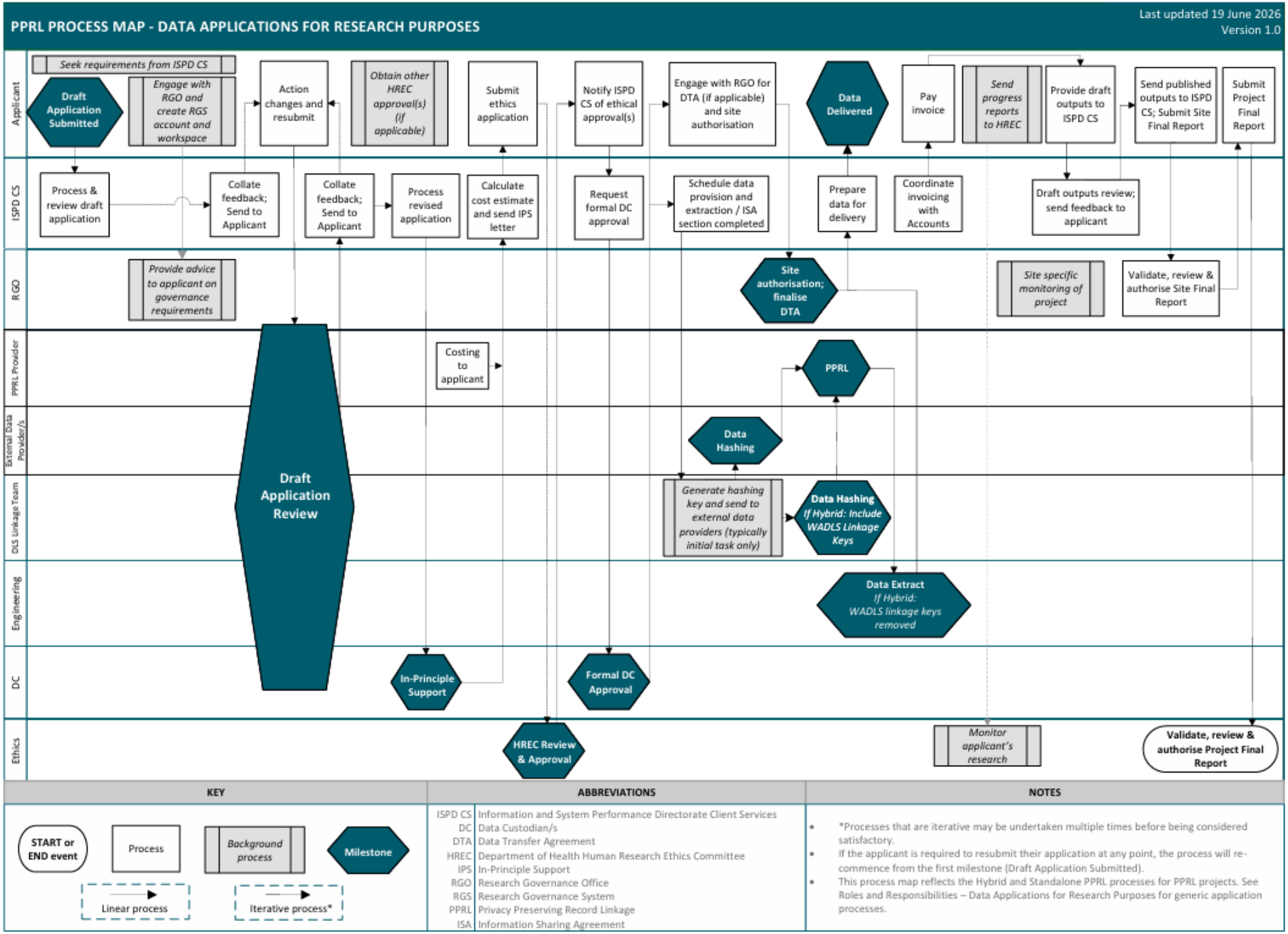
Acronyms	
BAU	Business-as-usual
CDL	Centre for Data Linkage
DTA	Data Transfer Agreement
HREC	Human Research Ethics Committee
ICWA	Insurance Commission of Western Australia
IPG	Information Policy and Governance
IRA	Information Release Agreement
ISA	Information Sharing Agreement
ISPD	Information and System Performance Directorate
LLS	Legal and Legislative Services
MOU	Memorandum of Understanding
OMRI	Office for Medical Research and Innovation
PPRL	Privacy Preserving Record Linkage
POC	Proof-of-concept
RGO	Research Governance Office
RGS	Research Governance Service
SOP	Standard Operating Procedure

Definitions	
Bloom filter	Bloom filters are used in PPRL to encrypt personally identifying information in a way that allows for comparisons of similarly spelled but not identical words. In PPRL, consecutive pairs of letters or numbers are encrypted. The Bloom filter is a space-efficient probabilistic data structure that is used to encode these hashed pairs into a single binary vector of fixed length. It is a hardening method involving modifications of bit-arrays representing identifiers to make attacks more difficult. The use of Bloom filters allows it to be possible to make matches between identifiers with small errors such as spelling mistakes.
Cohort	A collection of people who share an experience or characteristic over time and is often applied as a method of defining a population for the purpose of research.

Data	The term ‘data’ generally refers to unprocessed numbers, facts or statistics, while the term ‘information’ refers to data that has been processed in such a way as to be meaningful to the person who receives it. The terms ‘data’ and ‘information’ are often used interchangeably and should be taken to mean both data and information.
External Data Providers	Data providers located external to the Department of Health.
Hashed Data	In cryptography, hashing is a process that allows you to take data of any size and apply a mathematical process to it that creates an output that’s a unique string of characters and numbers of the same length. The hashing process is virtually irreversible, which means that hashing is a one-way cryptographic function.
Linkage Map	A linkage map is a file that contains linkage keys mapped to primary keys/record IDs and can be used for data integration.
Personally Identifiable Information	Includes information connected with an individual’s name, image, date of birth or address; information that contains a unique personal identifier when the holder of the information also has the master list linking the identifiers to individuals; and information that the holder can merge or link to other information they already hold, enabling them to identify individuals.
Personal Information	Defined in the *Freedom of Information Act 1992* (WA) as information or opinion about an identified individual.
Privacy Preserving Record Linkage (PPRL)	A method for linking sensitive databases across organisations without the need to reveal any private or confidential information. The most popular version of binary string encodings for PPRL are Bloom filters. The attractive main property of Bloom filters is that they can be used to encode strings in a similarity-preserving way.
Separation principle	The separation principle is a mechanism to protect the identities of individuals and organisations in datasets, applied as part of the data linkage and data integration process used to form the integrated dataset. It ensures that no individual can access identifying information and content data from multiple data sources simultaneously. This is achieved by assigning distinct roles across the linkage, extraction, integration and research processes, with each role permitted to access only the data necessary for its function. Under the separation principle, for example, data custodians can only view the datasets they are responsible for; linkers who require identifying information to create links cannot access substantive content data; and researchers or data users cannot access identifying information from any dataset.

8. Appendices

Appendix 1: PPRL Process Map



Appendix 2: External Data Providers checklist

Government of **Western Australia**
Department of **Health**

Privacy Preserving Record Linkage (PPRL)

EXTERNAL DATA PROVIDER CHECKLIST		**Last updated: 12 July 2023**
Instructions: Please fill the below checklist with information on the data you are providing		
Purpose: The quality and feasibility of performing linkage of data is dependent on the quality of data received from data providers. Hence, this checklist is developed to ensure minimum requirements are met to facilitate linkage.		
Element	**Format or Example**	**Value**
Date range (based on event date)	DD/MM/YYYY - DD/MM/YYYY	
Age range	XX - XX	
Specific Codes list or range (if applicable)	ICD Codes: A00 - B99, E00, L10 - L14 Pathology Codes: FBE, blood film, ECU, LFT, LDH, B12, Folate, TSHAIF, PCA, EPG, FLC, HaptoglobinBone marrow biopsy FBE, IS, ECU, LFT FOB, EPG, Hb-electrophoresisBlood Lead FBE, Reticulocyte count, ESR, ECU, CRPIS, B12, folate, LFTEPG, Bone marrow biopsy	
Sex/Gender distribution	Male = [n], Female = [n], Unknown = [n]	
Number of records/people provided	Records: [n] People: [n]	
Date of participant consent range (if applicable)	DD/MM/YYYY - DD/MM/YYYY	
Completion Date: _____		Printed Name: _____ Organisation: _____ Signature: _____

9. Document Control

This is a managed document. Changes should be issued as a complete replacement document with amendments recorded in the revision notes. Superseded versions should be removed from circulation.

This document will be reviewed every two years (or more frequently as required) and is amended via endorsement of the Data Linkage Strategy and Reform Program Control Group.

Document information

Title:	Privacy Preserving Record Linkage Guide
Owner:	Data Linkage, Data and Information Systems
Directorate:	Information and System Performance Directorate
TRIM reference:	D-AA-24/212045
Version:	2.0
Date of last review:	19/06/2026

Version history

Version	Date	Comments and notes	Approved by
0.1	06/06/2023	First draft.	
0.2	03/07/2023	Updated draft with changes requested by PPRL Working Group.	
1.0	17/08/2023	First version as endorsed by Data Linkage Executive Council.	Data Linkage Executive Council
1.1	25/05/2026	Updated version following joint review of guide by Data Linkage Services and CDL teams.	
2.0	19/06/2026	Version 2.0 endorsed by the Data Linkage Strategy and Reform Program Control Group	Data Linkage Strategy and Reform Program Control Group

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

© Department of Health 2026

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

health.wa.gov.au