

Data book
November 2025

Purpose of this document

This document provides:

- a snapshot of the data available in Lumos, and
- important considerations for analysis.

This document is intended to indicate to potential analysts whether Lumos is the appropriate tool for proposed work. After consulting this guide, applicants are encouraged to contact Lumos@health.nsw.gov.au for tailored information and to refine project applications.

Full details of Lumos datasets are available from, collectively, the Lumos Data Dictionary (available on request) and data dictionaries published on the Centre for Health Record Linkage (CHeReL) [website](#). Further technical information and guidance is available on the Lumos Community Sharepoint (available to approved Lumos users).

Information on Lumos for general practitioners and consumers, including security and privacy protections, is provided on the Lumos website and not in this document.

Lumos data are subject to change. Accordingly, this document should be considered a starting point for discussion only.

Version History

Version	Changes
November 2025	First document version

© NSW Ministry of Health

SHPN (HSSG) 250819

ISBN 978-1-74231-236-1

Contents

Purpose of this document	2
Version History	2
1-What is Lumos?.....	5
Datasets in Lumos.....	5
Lumos cohort and ‘linked-to’ datasets.....	5
2-Conditions of use.....	7
Lumos for planning, funding, management or evaluation of health services.....	7
Lumos for research purposes	7
Aims of Lumos	7
User responsibilities	8
3-Important notes for analysts.....	9
Aboriginal data	9
Data update frequency.....	9
Data sources.....	9
Privacy	9
Data availability and timeliness	9
Lumos coverage and representativeness.....	10
4 –Current data sets.....	12
General Practice records	12
Sources and Transformation	12
Variables	12
Healthdirect VirtualGP and nurse triage	14
Variables	14
NSW Health Urgent Care Services.....	16
LHD UCS	16
GP UCS	17
Healthdirect referrals to UCS	17
Ambulance ECP UCS.....	18
BreastScreen NSW.....	20
Variables	20
Mortality	22
Cause of Death	22
Death registrations	22
NSW Ambulance	23
Computer-Aided Despatch.....	23

Electronic Medical Record	23
Patient Health Care Record	24
NSW Cancer Registry.....	25
Variables	25
NSW Health Admitted Patients	26
Variables	26
NSW Health District and Network Returns	28
NSW Health Elective Surgery Wait times.....	29
Variables	29
NSW Health Emergency Department.....	31
Variables	31
NSW Health Integrating Care Database	33
Variables	33
NSW Health Mental Health Ambulatory	34
Variables	34
NSW Health Non-Admitted Patients (Outpatients).....	36
Variables	36
NSW Health Patient Reported Measures*	38
*Availability	38
Assessment tools	38
One-off patient cohorts	40
Purpose	40
Rules.....	40
5 - Future datasets.....	43
MBS.....	43
Purpose of including MBS data.....	43
Variables	43
NSW Health Waiting For What.....	44
Variables	44
Glossary	45

1 - What is Lumos?

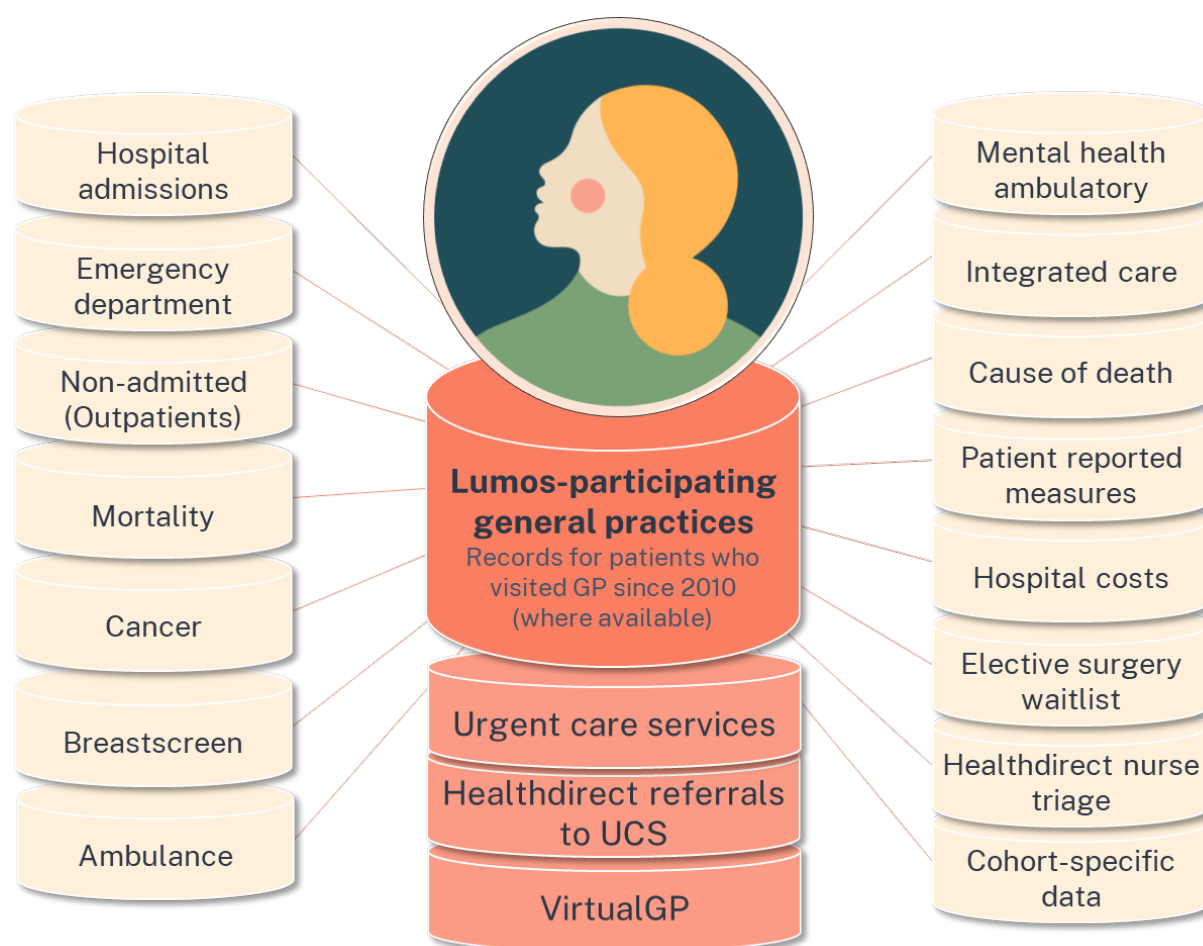
Lumos is an enduring linked data asset linking deidentified unit record patient data from participating NSW general practices to data held by NSW Health system and records from HealthDirect. This enables the following of patient journeys across the health system in NSW to inform system management improvements.

For further information on how Lumos fits with other data programs, see: [Lumos in the Health Data Landscape](#).

Data are available to analysts in a secure access environment called the SAPHE (Secure Analytics Primary Health Environment), which supports use of R, SQL and Python.

Datasets in Lumos

Lumos comprises the following datasets:



This list is subject to change.

Lumos cohort and 'linked-to' datasets

The following datasets comprise the Lumos cohort:

- Patients with General Practice Electronic Hospital Records

- Patients using a NSW Urgent Care Service¹
(NSW Health-funded GP-led clinics, and NSW Health-run urgent care services including the Extended Care Paramedics model)
- Patients with a Healthdirect referral to a NSW Urgent Care Service
- Patients referred to the Healthdirect VirtualGP service

Patient records in the cohort are ‘linked to’ the following datasets:

- Healthdirect VirtualGP and nurse triage²
- Mortality (cause of death and register of deaths)
- NSW Health datasets:
 - NSW Ambulance
 - NSW Cancer Registry
 - Admitted Patients
 - BreastScreen NSW
 - District and Network Returns (hospital costs)
 - Elective Surgery Wait Times
 - Emergency Department
 - Integrating Care Database
 - Non-Admitted Patients (outpatients)
 - Patient Reported Measures (not included in every linkage – see Section 4)
- One-off patient cohorts (see Section 4 for explanation)

The Lumos ethics approval includes the following additional datasets, which will be added at a later date:

- Medicare Benefits Schedule (MBS)
- NSW Health Waiting For What (Delays in patient discharge while waiting for discharge supports to be arranged)

What does this mean?

As the Lumos cohort is centred primarily around general practice records, the following logic applies:

- Where a patient attends any Lumos-participating general practice and a NSW Health Emergency Department, both their general practice and ED records are included in the Lumos data asset, **but**
- Where a patient attends a NSW Health Emergency Department but **does not** attend a Lumos-participating general practice, they will **not** be included in Lumos and their ED records are **not** included.

¹ Australian Government-run Urgent Care Centres are **not** in scope

² All patients in the Lumos cohort, *whether or not they had a referral to the Healthdirect VirtualGP service*, are linked to Healthdirect nurse triage helpline data where available

2 - Conditions of use

Lumos operates under ethical approval³ for the **planning, funding, management or evaluation of health services**. All uses must align with this approval.

Lumos use is regulated by the Lumos Data Governance Committee.

Lumos for planning, funding, management or evaluation of health services

The following rules apply:

1. Applicants not employed by a NSW Primary Health Network (PHN) or NSW Health entity **must be sponsored** by at least one NSW PHN or NSW Health entity. This is to ensure the work has a translational purpose to support the management of health services. Research with no translational element will not be permitted. Analysts from NSW Health and NSW PHNs may use Lumos data within the approved purposes without additional sponsorship.
2. Investigations must have well defined aims that align to the stated purposes of the Lumos program. **Undirected exploratory analyses by third parties will not be permitted.**
3. Analysis **must align** with at least one of the **Aims** of Lumos (see below).
4. Analysis should be intended to **benefit the health system in NSW** in whole or in part. Further, analyses should not be used for service performance management or benchmarking.
5. A **rationale for using Lumos** (including general practice data) is required.
6. **Lumos must be appropriate** for the planned analysis. Applicants are strongly encouraged to approach the Lumos team at lumos@health.nsw.gov.au to discuss planned analyses for feasibility.
7. Applicants must **consult all relevant stakeholders** prior to application.
8. Applicants must be **suitably qualified** to work with linked unit record health data using R, SQL or Python.

Lumos for research purposes

Per NSW legislation, ‘**Planning, funding, management or evaluation of health services**’ is distinct from ‘**Research**’. Lumos does not have ethics approval for research purposes.

However, third party applicants such as applicants from academia may apply to use Lumos provided the above rules are met. All applications will be reviewed by the Lumos Data Governance Committee.

Aims of Lumos

1. Provide information that will support public health programs in NSW in areas like chronic condition management, smoking, obesity, risky alcohol use, youth suicide and early intervention.
2. Improve the ability to target high risk groups in the population by determining reliable and modifiable predictors of poor health outcomes.

³ NSW Population & Health Services Research Ethics Committee ref. 2019/ETH00660

3. Discover the most effective leverage points to refine health service processes and to improve quality in patient care while reducing waste for a more cost-effective health system with better patient outcomes.
4. Improve identification of, and reductions in, preventable emergency department demand, unplanned hospital presentations and readmission rates.
5. Implement data-informed and evidence driven quality improvement, and system re-design to deliver better patient care.
6. Proactively monitor new service models to address emerging population health issues.
7. Foster greater collaboration between Local Health Districts (LHDs), PHNs and GPs to ensure the right care in the right place at the right time.

Further information is provided in the Lumos [Data Governance Framework](#).

User responsibilities

Due to the highly sensitive nature of the Lumos data asset, all SAPHE users are bound by the conditions of data release, terms of use and confidentiality undertaking for the Lumos data asset. This includes the following terms of use.

Users of the Lumos data through the SAPHE:

- Must not link or otherwise match Lumos data with information on individuals in other datasets
- Must not export any data that could expose identifiable patient-level information or clinical coding reference data
- Must not share access to, or disclose, Lumos data or personal information derived from the Lumos data in an identifiable form (e.g. through sharing log-ins or exporting/publishing/presenting data with small cell sizes)
- Must only use the data for the specific purposes or projects approved in their application and in accordance with the approved Lumos ethics protocol and all subsequent amendments
- Must provide the Lumos team with advanced notice of any planned publications
- May be subject to criminal penalties for misusing Lumos data.

When conducting analyses, users should consider not only the objectives of proposed analyses but any ethical implications if the results were to be published or insights actioned or implemented, including the potential impact(s) on vulnerable populations.

3 - Important notes for analysts

Aboriginal data

- Currently, there are **no Aboriginal health services** enrolled in Lumos, and existing data **cannot be segmented for Aboriginality**. Lumos is working with the Aboriginal Health and Medical Research Council (AH&MRC) and Aboriginal health leaders to embed Indigenous Data Sovereignty and Governance (IDS&G) into Lumos to address this matter.

Data update frequency

- Lumos data are **refreshed twice yearly in their entirety**. Each refresh is a new linkage. This is to ensure data reflect current consent. After refresh, no analysis with previous tranches is possible. Users are provided prior notice to export and/or update code as required.

See also: [Lumos Data Timing Explained](#); SAPHE Refresh FAQs and Lumos Data – What to be aware of (Lumos SharePoint)

Data sources

- Lumos is a **longitudinal data asset**; data reach back to 2010 where available.
- On July 2024, the official NSW Health data warehouse switched from the Health Information Exchange (HIE) to the Enterprise Data Warehouse (EDWARD). Accordingly, Lumos contains at least Admitted Patient, ED and Non-Admitted Patient data **from both HIE and EDWARD**.

See also: *Lumos Tranche 11 SAPHE release notes and HIE-EDWARD mapping (Lumos SharePoint)*

- General practice data are sourced from multiple systems via **different extraction vendors**, with no standardisation prior to receipt by Lumos (see Section 4).

Privacy

- All data in Lumos are **deidentified** in accordance with RACGP⁴ and AMA⁵ expectations. Lumos uses probabilistic Privacy Preserving Record Linkage (PPRL) underpinned by bloom filtered encoding of patient identifiers using LinXmart. Linkage accuracy is 96-99.3% compared to clear text linkage^{6,7}. As the linkage rate is not 100%, Lumos is **not a clinical resource**.

See also: [Lumos in the Health Data Landscape](#)

- Lumos includes **only structured data** extracted from health services. **Free text variables are excluded** on privacy grounds.
- Unit record data cannot be exported from the SAPHE**. Export requests will be vetted for sufficient aggregation to prevent reidentification.

Data availability and timeliness

- Data are **segmented by PHN**. Access is typically provided for the relevant PHN for each project; **multi-region access requires additional approval**.

⁴ RACGP Three key principles for the secondary use of general practice data by third parties

⁵ AMA Position Statement on Data Governance and Patient Privacy in Healthcare

⁶ Irvine, K. et al. Real world performance of privacy preserving record linkage. International Journal of Population Data Science., 3 (4) (2018)

⁷ Randall, S. et al. A blinded evaluation of privacy preserving record linkage with Bloom filters BMC Medical Research Methodology, (2022) 22:22

- Data sources and variables included in Lumos may **change over time** due to changes in the Lumos ethics approval, data availability, and frequency of use. All users are advised of changes via the SAPHE Users Community of Practice.
- Lumos data are sourced from multiple collections with different availabilities and lags. Users are advised of **different dataset date ranges** with each data refresh.
- Lumos data are **made available approximately 6 months after extraction from general practices**. This lag must be factored into analytics timelines.

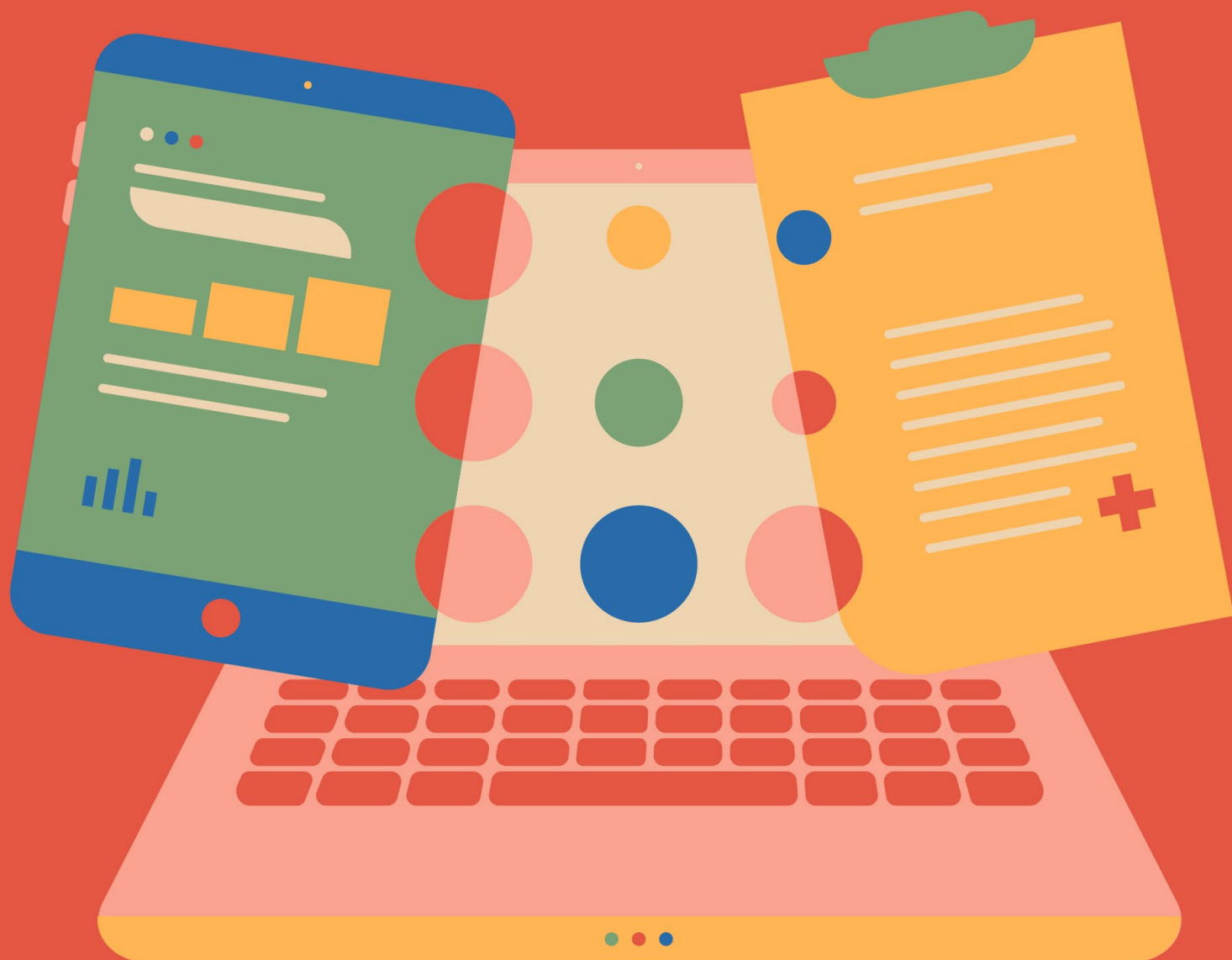
See also: [Lumos Data Timelines Explained](#)

Lumos coverage and representativeness

- General practices opt in to participate in Lumos and the number of participating practices continually changes and grows over time. As of November 2025, Lumos included over **890 general practices (40% of NSW general practices)** and **7 million patients** (approximately **50-60% of the NSW population** in any given year).
- Accordingly:
 - Lumos is a **sample of the NSW population**; appropriate adjustments may be required during analysis.
See also: *Lumos Data – What to be aware of (Lumos SharePoint)*
 - Lumos **does not capture any information about patients who only attend a non-participating general practice**.
 - It is currently not possible to determine the extent to which a Lumos patient within Lumos also receives care at a non-participating general practice. That is, the **full coverage of patients' general practice health journeys is unknown**. This will be addressed with the addition of MBS data on general practices claim events (see Section 5).
- Independent analyses indicate Lumos is strongly representative of the NSW population in relation to sex, rurality and age, geography, and most socioeconomic measures⁸.

See also: [Lumos Representativeness Factsheet](#)

⁸ Bouckley T, Myton-Katieva R, Peiris D, et al. An assessment of data quality and sociodemographic variation in health service utilisation of general practice, emergency department and admitted services in a New South Wales linked health data asset: a retrospective cohort study of Lumos. *BMJ Open* 2025;15:e102055.



4 – Current data sets



General Practice records

Name of dataset	General Practice Electronic Health Records (GPEHR)
Source	Individual participating NSW general practices via extraction vendors (see notes below)
Linkage dataset type	Cohort See Section 1 for further information on Cohort datasets
See also	Lumos user guides (Lumos Sharepoint) GP data in Lumos (Lumos Sharepoint)

Sources and Transformation

There are currently three vendors which provide GP data to Lumos: PenCS, Outcome Health and PrimarySense. Each extraction vendor's software collates and harmonises data from across several different GP clinical information systems (Best Practice, Medical Director, etc.). Completeness and harmonisation of data differ between extraction vendors.

With the exception of Central and Eastern Sydney PHN (CESPHN), generally each PHN uses a single vendor from these options. Users who have access to data from multiple PHNs, or from CESPHN, are advised that many differences arise in the Lumos GP data as a result of use of different vendors and their bespoke data transformation.

Variables

This list presents the variables approved for inclusion in Lumos. Not all variables may be available for each tranche. Analysts are advised to discuss analytics needs with the Lumos team prior to application.

For full details for variables listed below, see the Lumos Data Dictionary (available on request and on the Lumos SharePoint).

Variable Name	Notes
Clinical information system of source data	
Practice ID	Identity of participating general practices is not made available to external analysts
Active patient flag	Identifies patients who have attended the practice 3 or more times in the preceding 2 years
Patient status	E.g. active, archived or deceased
Age in years at date of extract	
Sex	
DVA status	

Marital Status	
Encounter counts and dates	
Provider type	GP, nurse, etc.
MBS Item number, date and description	
Chronic disease flags, diagnoses and dates	Includes: diabetes, COPD, asthma, heart failure, coronary heart disease, mental health conditions (anxiety, bipolar, schizophrenia, depression, autism, ADHD, dementia), chronic kidney disease, osteoarthritis
Program cohort: <ul style="list-style-type: none"> • flags and sub-cohort flags, • mode of entry, • condition type, • procedure (if any), • immediate outcome, • disability status (if any), • alternative options for patient care, • language spoken at home, • interpreter required, • whether a patient has a regular GP, • redirection to other care, • on-referrals (if any), • mode of clinical handover 	For patients participating in specific programs, e.g. urgent care. Applicants seeking to evaluate a program are recommended to discuss whether a program cohort flag can be made available to support analysis with the Lumos team.
Medications prescribed over time	Includes generic name, first date, last date, reason for prescription and number of repeats. Does not include dispensing data.
Lifestyle health risk factors and dates	Includes smoking status and alcohol use
Physical parameters and dates	Includes blood pressure, spirometry, height and weight
Biomedical measurements and dates	Includes HBA1c, cholesterol, low density lipoprotein, high density lipoprotein, triglycerides, eGFR, microalbumin, liver function tests, haemoglobin and blood glucose
Health screening and dates	
Immunisation and dates	Provided immunisations are limited by the extraction vendor, with PenCS providing only pneumococcal and influenza vaccinations.
My Health Record flag	
Dates of shared health summary uploads	
MyMedicare flag	Not currently available
Reason for visit	Not currently available
SEIFA, MMM	Derived from other data — presented in summary tables, not in raw GPEHR tables

Healthdirect VirtualGP and nurse triage

Name of dataset	Healthdirect VirtualGP and nurse triage
Source	Healthdirect
Linkage dataset type	Cohort: all patients referred by Healthdirect nurse triage to VirtualGP service Linked-to: nurse triage records (where available) for all patients in cohort See Section 1 for further information on Cohort and Linked-to datasets

Variables

This list presents the variables approved for inclusion in Lumos. Not all variables may be available for each tranche. Analysts are advised to discuss analytics needs with the Lumos team prior to application.

For full details for variables listed below, see the Lumos Data Dictionary (available on request and on the Lumos SharePoint).

Variable	Notes
Nurse Triage	
Service Provided	
Session Start and End Time	
Call Type and Sub-type	e.g. Emergent Call, Triage Call
Caller: <ul style="list-style-type: none"> Gender State Relationship to patient 	
Patient: <ul style="list-style-type: none"> Gender Age Ethnicity Preferred language Cultural background 	
Patient Consent	Whether a Patient is eligible and has given Consent to be transferred to a GP under the GP After Hours service
Triage Final Outcome	Referral end point
Triage Inclination Description	The medical action the Patient would have taken if it were not for the Triage Service
Initial symptoms	
Risk factors	
VirtualGP	
Attempt Counter	The requirement for a second outbound call attempt when the initial attempt was not successful.
Age	
Callback and Handling Time	
Gender	

GP Outcome	The outcome as recorded by the GP following VirtualGP Service.
Patient Preferred Language	
Tier	Priority level for GP callbacks based on Triage Nurses' assessments
Video Call	
GP Decision Support Tool	
Health Summary Consent	Patient consent to share health summary

NSW Health Urgent Care Services

The lists in this section present the variables approved for inclusion in Lumos. Not all variables may be available for each tranche. Analysts are advised to discuss analytics needs with the Lumos team prior to application.



LHD UCS

Name of dataset	NAP Data Collection LHD UCS cohort – subset of NAP
Source	NSW Health
Linkage dataset type	Cohort See Section 1 for further information on Cohort datasets
See also	https://www.health.nsw.gov.au/Hospitals/Pages/urgent-care-services.aspx CHeReL website .

Highlighted variables indicate differences between the variables available in Lumos for UCS patients (here) versus other patients in the Non-Admitted Patient dataset (see NSW Non-Admitted Patients, below).

Variable	Notes
Service Unit Full Name	
Service Unit HERO Identifier	
UCS mode of entry	Includes walk-in, local ED, health info and triage website or phone line, ambulance diversion, GP, other health professional, other
UCS condition type	Includes acute wound, injury, acute illness, acute exacerbation of known chronic condition, routine management of known condition, other.
UCS procedure (if any)	Includes: intravenous therapy, sutures, plastering, eyecare, wound dressing, incision, abscess drainage, catheter management
Immediate outcome	Includes: home, hospital, ED, GP referral, did not wait, other
Disability status (if any)	Includes sensory, intellectual, physical, psychological, head injury, stroke, other.
Alternative options for patient care	Local ED, health info and triage website or phoneline, ambulance, GP, other health professional, other
Language spoken at home	
Interpreter required	
Has a regular GP	
Redirection to other care (if any)	
On-referrals (if any)	
Mode of clinical handover	



GP UCS

Name of dataset	General Practice Electronic Health Records UCS module – bespoke subset of GPEHR
Source	Individual participating NSW general practice-led UCSs via extraction vendors
Linkage dataset type	Cohort See Section 1 for further information on Cohort datasets
See also	https://www.health.nsw.gov.au/Hospitals/Pages/urgent-care-services.aspx CHeReL website .

Refer to General Practice records (Program Cohort variables).



Healthdirect referrals to UCS

Name of dataset	Healthdirect referrals to NSW Health UCS
Source	NSW Health (EDWARD)
Linkage dataset type	Cohort See Section 1 for further information on Cohort datasets
See also	https://www.health.nsw.gov.au/Hospitals/Pages/urgent-care-services.aspx Lumos Data Dictionary

Variable	Notes
Referral date and time	
Interpreter	
Gender	
Age	
Home LHD	
Symptoms	
Presentation category	
Triage	
Referred to service name	
Referred to service appointment time (if applicable)	
Referred to service appointment date	



Ambulance ECP UCS

Computer Aided Despatch:

Name of dataset	NSW Ambulance Computer-Aided Despatch (CAD) Extended Care Paramedics (ECP) UCS – subset of CAD
Source	NSW Ambulance
Linkage dataset type	Cohort See Section 1 for further information on Cohort datasets
See also	https://www.ambulance.nsw.gov.au/news/news-items/paramedics-join-tweed-urgent-care-service-in-a-nsw-first

Highlighted variables indicate differences between the variables available in Lumos for ECP patients (here) versus other patients using Ambulance services (see NSW Ambulance, below).

Variable	Notes
Response ID	
Response_date	
Zones	
Pick up location	
Priority number	
Response time	
City	
State	
Matrix allocation codes	Paramedic-initiated query to determine nearest most suitable hospital
Problem	
Incident Cancel Reason	

Ambulance Electronic Medical Record:

Name of dataset	NSW Ambulance Electronic Medical Record (eMR) Extended Care Paramedics (ECP) UCS – subset of Ambulance eMR
Source	NSW Ambulance
Visibility in Lumos	Cohort
See also	https://www.ambulance.nsw.gov.au/news/news-items/paramedics-join-tweed-urgent-care-service-in-a-nsw-first

Highlighted variables indicate differences between eMR variable availability in Lumos for ECP patients versus other patients using Ambulance services.

Variable	Notes
Fleet unit	
Case/incident number	

Case Date	
Case description	
Dispatch code	
Scene (location name)	
Scene (State)	
Patient Gender	
Patient Age	
Patient address (State)	
Not transported reason	
Not treated reason	
Risk Factors	
Case Nature	
Billing type	
Referral	
Observed patient outcome	

BreastScreen NSW

Name of dataset	Breastscreen NSW
Source	Cancer Institute NSW
Linkage dataset type	Linked-to See Section 1 for further information on Linked-to datasets

Variables

This list presents the variables approved for inclusion in Lumos. Not all variables may be available for each tranche. Analysts are advised to discuss analytics needs with the Lumos team prior to application.

For full details for variables listed below, see the CHeReL [website](#).

Variable	Notes
Client	
Birth year, month	
Death year, month	
Cause of death	This variable may be subject to significant data lag
Main language other than English spoken at home	
Family and previous history of breast cancer	
Episode	
Episode Number	
Date of First Attendance	
Screening Visit	
Attendance Date	
Screening Recommendation	
Screening Visit Symptom	
Symptom Status and Laterality	
Assessment Visit	
Attendance Date	
Assessment Recommendation	
Percutaneous Needle Biopsy Result – Cytology and Histology	
Excision performed Y/N, date and result	
Recommendation and date of definitive diagnosis	
Histopathology of Malignant Lesions	
Interval Cancer	
Date of Detection	
Histopathology of Malignant Lesions	
Size of tumour	
Histological grade	
Date of commencement of treatment	
Surgical treatment	
Radiotherapy	
Chemotherapy	
Metastasis – distant	

Mortality

The lists in this section present the variables approved for inclusion in Lumos. Not all variables may be available for each tranche. Analysts are advised to discuss analytics needs with the Lumos team prior to application.

For full details for variables listed below, see the CHeReL [website](#).

Cause of Death

Name of dataset	Cause of Death (COD)
Source	Australian Coordinating Registry
Linkage dataset type	Linked-to See Section 1 for further information on Linked-to datasets

To meet audit requirements, **all Lumos applicants must provide a valid justification for access to cause of death data at the time of application.**

Variable	Notes
Year of registration	
Date of death	
Age at death	
State/Territory of Usual residence	
Place of Death	
Underlying Cause of Death Diagnosis Code	
Contributing causes of death (ICD-10)	
ABS revision status	To adjust for reporting delay

Death registrations

Name of dataset	Death registrations
Source	Registry of Births, Deaths and Marriages (RBDM)
Linkage dataset type	Linked-to See Section 1 for further information on Linked-to datasets

Variable	Notes
Date of death	
Age at death in years	

NSW Ambulance

The lists in this section present the variables approved for inclusion in Lumos. Not all variables may be available for each tranche. Analysts are advised to discuss analytics needs with the Lumos team prior to application.

For full details for variables listed below, see the CHeReL [website](#).



Computer-Aided Dispatch

Name of dataset	NSW Ambulance Computer-Aided Dispatch (CAD)
Source	NSW Ambulance
Linkage dataset type	Linked-to See Section 1 for further information on Linked-to datasets

Variable	Notes
Response date	
Priority number	
Response time	
City	
State	
Matrix allocation codes	Paramedic-initiated query to determine nearest most suitable hospital
Problem	



Electronic Medical Record

Name of dataset	NSW Ambulance Electronic Medical Record (eMR)
Source	NSW Ambulance
Linkage dataset type	Linked-to See Section 1 for further information on Linked-to datasets

Variable	Notes
Case Date	
Case description	
Dispatch code	
Scene: <ul style="list-style-type: none"> • Location • State 	
Scene (State)	
Patient: <ul style="list-style-type: none"> • Gender • Age 	
Not transported reason	

Not treated reason	
Risk Factors	
Case Nature	
Billing type	
Referral	
Observed patient outcome	



Patient Health Care Record

Name of dataset	NSW Ambulance Patient Health Care Record (PHCR)
Source	NSW Ambulance
Linkage dataset type	Linked-to See Section 1 for further information on Linked-to datasets

Variable	Notes
Transport Date	
Master Incident ID	
Source	
Priority Code Description	
Description of Dispatch Code	e.g. 2 - Acute Non Time Critical
Patient: <ul style="list-style-type: none"> Sex Age 	
Disposition Code and Description	e.g. Transported, non-transport
From: <ul style="list-style-type: none"> Hospital Code/Name/Type Suburb 	Refers to location to which ambulance is despatched
To: <ul style="list-style-type: none"> Hospital Code/Name/Type 	
Condition	
Date of Incident	
Problem Code/Description	
Referral Attempted	
Referral Type	
Referral Outcome	

NSW Cancer Registry

Name of dataset	NSW Cancer Registry (CCR)
Source	Cancer Institute NSW
Linkage dataset type	Linked-to See Section 1 for further information on Linked-to datasets

Variables

This list presents the variables approved for inclusion in Lumos. Not all variables may be available for each tranche. Analysts are advised to discuss analytics needs with the Lumos team prior to application.

For full details for variables listed below, see the CHeReL [website](#).

Variable	Notes
Sex	
Country of Birth	
Birth year, month	
Death: <ul style="list-style-type: none"> • year, month • ICD-10-AM Cause of death • Place of death 	
Diagnosis: <ul style="list-style-type: none"> • Date of Diagnosis • Age at Diagnosis • Cancer Group • Clinical Group • Best basis of diagnosis 	
ARIA, LGA, LHD, PHN, SEIFA, SLA	



NSW Health Admitted Patients

Name of dataset	Admitted Patient Data Collection (APDC)
Source	NSW Health EDWARD: from 1 July 2024 HIE: 2010 to 30 June 2024
Linkage dataset type	Linked-to See Section 1 for further information on Linked-to datasets

Variables

This list presents the variables approved for inclusion in Lumos. Not all variables may be available for each tranche. Analysts are advised to discuss analytics needs with the Lumos team prior to application.

For full details for variables listed below, see the CHeReL [website](#).

Variable	Notes
Hospital: <ul style="list-style-type: none"> Hospital type (public/private) Acute Hospital flag LHD Facility transferred from/to, type Peer group 	Linkage for private hospitals may be less accurate than for public. Hospital type variable indicates the potential for false links. Peer group: Public hospitals only
Patient: <ul style="list-style-type: none"> Age and age group Birth date (year and month) Country of birth and usual residence DVA card type Health insurance on admission Marital status Payment status on separation Sex State of residence LGA, LHD, PHN, SLA Financial class, program and sub program 	
Episode/care: <ul style="list-style-type: none"> Start date, time End date, time LOS Care type Leave days total Episode Sequence number Australian Refined Diagnosis Related Group (AR-DRG) 	

<ul style="list-style-type: none">• Clinical codeset• Condition onset flags• Days in psychiatric unit• Diagnosis codes• DRG mode of separation• Emergency status• ED Status• Hours in ICU• Hours on mechanical ventilation• Involuntary days in psychiatric unit• Last psychiatric admission date• Major Diagnostic Category• Mode of separation• Procedure block number, codes, date and location• Bed days: Qualified bed days and Unqualified baby bed days• Referred to on separation• Source of referral• Service Related Group (SRG)• Stay number (encrypted)• Unit type on admission• NWAU	<p>Stay number is hashed (encrypted) before receipt by Lumos</p>
---	--

NSW Health District and Network Returns

Name of dataset	District and Network Returns (DNR)
Source	NSW Health
Linkage dataset type	Linked-to See Section 1 for further information on Linked-to datasets Note: data are appended to NSW Health Admitted Patient (APDC) records and not directly linked
Purpose	Indicates costs of hospital treatment

Further information on hospital cost data available on request.

NSW Health Elective Surgery Wait times

Name of dataset	Elective Surgery Wait Times (ESWT)
Source	NSW Health
Linkage dataset type	Linked-to See Section 1 for further information on Linked-to datasets

Variables

This list presents the variables approved for inclusion in Lumos. Not all variables may be available for each tranche. Analysts are advised to discuss analytics needs with the Lumos team prior to application.

Variable	Notes
Facility Code	
Accommodation: <ul style="list-style-type: none"> Status, Dates of accommodation status changes Elected on admission 	
Patient listing: <ul style="list-style-type: none"> Listing date Transaction Date of Listing Date Status Times listing status changed Dates of listing status changes Days “ready” and “not ready for care” 	Status examples: 1 - Ready for Care, to 2 - Deferred
Admission dates: <ul style="list-style-type: none"> Planned Provisional Dates of admission date changes Number of times patient delayed on planned admission date 	
Procedure: <ul style="list-style-type: none"> Planned procedure date Intended LOS (Discharge Intention) Indicator procedure code Specialty code Current and Previous Urgency Categories; dates of change between categories 	
Delays: <ul style="list-style-type: none"> Status Times delayed in current period Total delays Dates of delays Number of days the patient has waited for the booked procedure 	Delays include: No theatre, no bed, no doctor, and not specified
Declines, removals and deferrals: <ul style="list-style-type: none"> Decline Status 	

<ul style="list-style-type: none">• Number of patient declined offers of admission• Date of last patient declined offer of admission• Reason for and date of removal from wait list• Number of deferrals	
Wait List Record Deletion Indicator	



NSW Health Emergency Department

Name of dataset	Emergency Department Data Collection
Source	NSW Health EDWARD: from 1 July 2024 HIE: 2010 to 30 June 2024
Linkage dataset type	Linked-to See Section 1 for further information on Linked-to datasets

Variables

This list presents the variables approved for inclusion in Lumos. Not all variables may be available for each tranche. Analysts are advised to discuss analytics needs with the Lumos team prior to application.

For full details for variables listed below, see the CHeReL [website](#).

Variable	Notes
Arrivals: <ul style="list-style-type: none"> Referral source Date Time Mode of arrival 	
Departures: <ul style="list-style-type: none"> Ready date and time Actual date and time Mode of separation Referred to 	
Triage: <ul style="list-style-type: none"> Category Date Time 	
Care: <ul style="list-style-type: none"> Principal ED diagnosis Doctor seen date, time Nurse practitioner seen date, time Clinical codeset Type of visit NWAU 	
Patient: <ul style="list-style-type: none"> Age and age group Birth date (year and month) Compensable status Country of birth DVA card type Marital status Need for interpreter 	

<ul style="list-style-type: none">• Preferred language• Sex• State of usual residence• LGA, LHD, PHN, SLA	
Facility: <ul style="list-style-type: none">• Facility• LHD• Facility type	

NSW Health Integrating Care Database

Name of dataset	NSW Health Integrating Care Database
Source	NSW Health
Linkage dataset type	Linked-to See Section 1 for further information on Linked-to datasets
Purpose	Flags patient enrolment in NSW Health Integrated Care programs

Variables

This list presents the variables approved for inclusion in Lumos. Not all variables may be available for each tranche. Analysts are advised to discuss analytics needs with the Lumos team prior to application.

Variable	Notes
Age at entry date	
Gender	
LHD	
Integrated Care program	Refer to https://www.health.nsw.gov.au/Value/ for more information
Integrated care enrolment date	
Previous Chronic Disease Management	
Sub-cohort	



NSW Health Mental Health Ambulatory

Name of dataset	NSW Mental Health Ambulatory Data Collection (MH-AMB)
Source	NSW Health
Linkage dataset type	Linked-to See Section 1 for further information on Linked-to datasets

Variables

This list presents the variables approved for inclusion in Lumos. Not all variables may be available for each tranche. Analysts are advised to discuss analytics needs with the Lumos team prior to application.

For full details for variables listed below, see the CHeReL [website](#).

Variable Name	Notes
Activity/Care: <ul style="list-style-type: none"> Start time Code Mental Health Codeset Duration Start date Group Identifier, Reason Mental Health Target Group Type Multidisciplinary Flag Principal Service Category Referred to Further Care Type Service Contact Mode Service Delivery Location Type Source of Referral Type Service Unit Identifier Facility identifier 	
Mental Health Diagnosis: <ul style="list-style-type: none"> Diagnosis group Diagnosis code Additional diagnosis 	Code: The mental health diagnosis code chiefly responsible for occasioning the client/patient's presentation to a community mental health care facility Additional diagnosis: up to 3 fields
Patient: <ul style="list-style-type: none"> Age Residence Country of birth DVA health card type Client Present Status Financial Class 	Residence: Area, health service, LHD, State

<ul style="list-style-type: none">• Legal Status• Marital Status• Preferred Language• Sex• Service recipient type	
Mental Health Provider Role, type	
Provider Award State, Financial Program	
Contact sequence number	



NSW Health Non-Admitted Patients (Outpatients)

Name of dataset	Non-Admitted Patient Data Collection (NAP)
Source	NSW Health EDWARD: from 1 Jan 2025 HIE: 2010 to 31 Dec 2024
Linkage dataset type	Linked-to See Section 1 for further information on Linked-to datasets
See also	Independent Health and Aged Care Pricing Authority (IHACPA) Tier 2 Non-Admitted Services Classification

Variables

This list presents the variables approved for inclusion in Lumos. Not all variables may be available for each tranche. Analysts are advised to discuss analytics needs with the Lumos team prior to application.

For full details for variables listed below, see the CHeReL [website](#).

Variable	Notes
Patient: <ul style="list-style-type: none"> Client age at time of service DVA Cover Sex ASGC Remoteness Category and Description State Country of residence and country of birth Preferred language Accommodation type code Usual Living Arrangements 	
Request Source Type	
Service Event <ul style="list-style-type: none"> Start date, time End date, time Client Participated Flag Group Session Flag Individual Service Provider Discipline Specialty Initial Or Subsequent Service HERO Identifier Non-admitted Financial Class NWAU Non-Admitted Patient Clinical or Therapeutic Service Flag Non-admitted Patient Service Type Non-Admitted Patient Service unit Establishment Type Primary Activity Type 	

<ul style="list-style-type: none">• Primary Setting Type• Responsible Organisation Service Provider• Responsible Service Facility and Provider Identifier• Service Contact Mode – Non-admitted Patient• Service Unit Full Name• TIER 2 Non-Admitted Service Classification	
---	--

NSW Health Patient Reported Measures*

Name of dataset	Patient reported measures surveys from the Health Outcomes and Patient Experience (HOPE) system
Source	NSW Health
Linkage dataset type	Linked-to; only included in Lumos on request See Section 1 for further information on Linked-to datasets
See also	https://aci.health.nsw.gov.au/statewide-programs/prms

*Availability

Patient Reported Measures are not included in every Lumos tranche. Please contact Lumos@health.nsw.gov.au if you require these data to ensure inclusion in a relevant linkage.

Assessment tools

It is not possible to fully list all variables for patient reported measures (PRMs), as each tool records different variables. Data from following tools are permitted for inclusion in Lumos (note inclusion is subject to availability):

- COVID-19 Yorkshire Rehabilitation Screening Tool
- Depression Anxiety Stress Scale-21
- Diabetes Distress Screening Scale
- Health Questionnaire (EQ-5D-5L)
- Falls Efficacy Scale (FES-I)
- Hip Dysfunction and Osteoarthritis Outcome Score (HOOS)
- Integrated Palliative Outcome Scale (IPOS) Renal
- Kansas City Cardiomyopathy Questionnaire (KCCQ)
- Knee Injury and Osteoarthritis Outcome (KOOS)
- Oxford Hip score
- Oxford Knee score
- Patient Reported Outcome Measures Information System (PROMIS-29)
- Post COVID-19 Functional Status (PCFS) Scale
- Problem Areas in Diabetes (PAID) Scale
- St George Respiratory (SGR) Questionnaire
- Wound-Q
- Emergency department PREM
- Inpatient PREMs
- Longitudinal PREMs
- Paediatric Integrated Care Survey
- Outpatient PREMs
- Paediatric PREMs
- Virtual care PREMs
- AQL-6D Questionnaire
- Cardiff Wound Impact Schedule (CWIS)
- Catquest-9SF
- COPD (Chronic Obstructive Pulmonary Disease) Assessment Test (CAT)

- International Consultation on Incontinence Questionnaire - Urinary Incontinence Short Form (ICIQ-UI-SF)
- Menopause-specific Quality of Life Questionnaire (MENQOL)
- mRS (modified Rankin Scale) simplified questionnaire
- Paediatric Quality of Life Inventory (PedsQL)
- Partners in Health Scale Self-Assessment (PIH)
- Patient Health Questionnaire-4 (PHQ-4)
- PROMIS Pediatric Profile v2.0 (PROMIS 25)
- PROMIS29+2
- PROMIS Profile CAT v1.0 (PROMIS CAT)
- Wexner Cleveland Clinic Incontinence Score (Wexner Scale)

One-off patient cohorts

Purpose

Lumos has ethical approval for one-off linkages of external patient cohorts as a Linked-to data collection to the Lumos Data Asset for the purposes of evaluating a program.

This is to enable investigation of health care delivery for a subset of people who participate in a particular program AND are already present in the Lumos dataset (i.e. it allows their participation in the program to be flagged within Lumos); if a participant has not visited a Lumos-participating practice or Urgent Care Service, their data will not be linked through the one-off linkage.

Rules

The following review criteria are used by the Data Governance Committee to assess an application to include a one-off cohort into Lumos.

For further information, please contact Lumos@health.nsw.gov.au.

1. **No new information** about program participants will be added to the Lumos Data Asset.
 - *Only the patients' participation in the program will be added to Lumos.*
2. **The program to be evaluated has a close connection to GPs**, such as a GP-delivered intervention, or a program that is expected or designed to impact on the delivery of primary care by GPs.
3. The program to be evaluated **has not been described to program participants as an 'anonymous'**, de-identified or similar service; and any program participants who sought to remain anonymous or pseudonymous while participating in the program or who opted-out from other use of their data, can be excluded from data linkage.
4. The objective of the evaluation relates to, or has implications for, the **funding, management, planning or evaluation of health** services in NSW, beyond just the particular program being evaluated.
5. The **results of the program evaluation will be shared** with the Ministry of Health for the benefit of the people of NSW.
6. Linkage of data relating to the program participants with data about the same patients as found in the Lumos Data Asset is **'reasonably necessary' for successful evaluation** of the program.
7. **Lumos data are 'fit for purpose'** for the program evaluation.
8. The program provider has **lawful authority** to disclose data about their program participants to NSW Ministry of Health (and reciprocally, the Ministry has lawful authority to collect data about the program participants), on the basis of either
 - a. the express consent of program participants, or

- b. a waiver of consent from a Health Research Ethics Committee which has approved the data linkage aspect of the evaluation on grounds relating to the funding, management, planning or evaluation of health services in NSW.
- 9. Identifying data from the program being evaluated **can be 'bloomed'** to implement privacy preserving record linkage (PPRL) technology to encode directly identifying particulars prior to linkage and integration in the CHeReL to create the one-off program evaluation dataset.
- 10. The entity conducting the program evaluation must be, or has **partnered** with, the Ministry of Health, a PHN or LHD such as through a memorandum of understanding or binding contract.
- 11. The one-off program evaluation dataset will be held and access controlled under the same conditions as the Lumos data asset (e.g. in a **separate part of the SAPHE**). Access will only be for the purpose of the one-off program evaluation project. The dataset will be destroyed as soon as possible after the program evaluation project is completed.



5 - Future datasets

The Lumos asset continually grows according to need. Feedback on possible inclusions is welcome via Lumos@health.nsw.gov.au. This section outlines datasets for which ethics approval has already been received, with implementation pending.

MBS

Name of dataset	Medicare Benefits Scheme (MBS)
Source	Australian Government Department of Health, Disability and Aged Care (DHDA)
Linkage dataset type	Planned Linked-to dataset See Section 1 for further information on Linked-to datasets
See also	MBS Online

Purpose of including MBS data

Lumos contains the general practice records and linked NSW Health records of patients who attend a participating general practice. A gap in Lumos therefore arises where patients:

- receive care in general practices that do not participate, or
- receive only part of their care in Lumos-participating general practices.

Lumos has ethics approval for the below MBS variables to address this gap. MBS data for patients in the Lumos cohort will indicate the extent to which care is provided by non-participating general practices, and therefore enable both a more comprehensive view of Lumos coverage and corrections in analyses.

Note: MBS variables as recorded at participating general practice are already included in Lumos per the General Practice dataset (see Section 4).

Variables

Variables as per Lumos ethics approval (high level):

Variable	Description/Notes
Date of Service	
Item Category	GP appointments to only include: standard, long, telehealth. Specific item numbers: GP attendance: all items in Groups A1, A2, A5, A6, A7, A11, A14, A15 Subgroup 1, A12 Subgroup 2 (items 735-758), A17, A18, A19, A20, A22, A23, A23, A30
Item Description	
MBS Item Number	
Schedule Fee	

NSW Health Waiting For What

Name of dataset	Waiting For What (W4W)
Source	NSW Health (EDWARD)
Linkage dataset type	Planned Linked-to dataset See Section 1 for further information on Linked-to datasets
See also	TBC

Variables

Variables as per Lumos ethics approval (high level):

Variable	Notes
Admission Date Time	
Area Abbreviation	
Care Type	
Discharge Date Time	
Facility	
Visit Identifier	
Waiting For What Category	
Waiting For What Close Date	
Waiting For What Created Date	
Waiting For What Effective Start Date	
Waiting For What Identifier	
Waiting For What Reason	

Glossary

AH&MRC	Aboriginal Health and Medical Research Council
AMA	Australian Medical Association
CHeReL	Centre for Health Record Linkage
DVA	Department of Veterans Affairs
ED	Emergency Department
EDWARD	Enterprise Data warehouse for Analysis, Reporting and Decision Support
HIE	Health Information Exchange
IDS&G	Indigenous Data Sovereignty and Governance
LHD	Local Health District
MBS	Medicare Benefits Schedule
NAP	Non-Admitted Patient
PHN	Primary Health Network
PPRL	Privacy Preserving Record Linkage
RACGP	Royal Australian College of General Practitioners
SAPHE	Secure Analytics Primary Health Environment
UCS	Urgent Care Service

