

Data linkage in Victoria

7 August 2017

Sharon Williams, Manager, Centre for Victorian Data Linkage

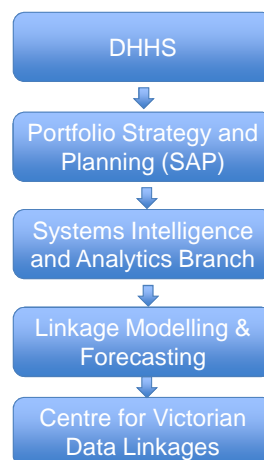
Mark Siphthorp, Assistant Manager, Centre for Victorian Data Linkage

Paul Schollum, Principal Analyst, Centre for Victorian Data Linkage

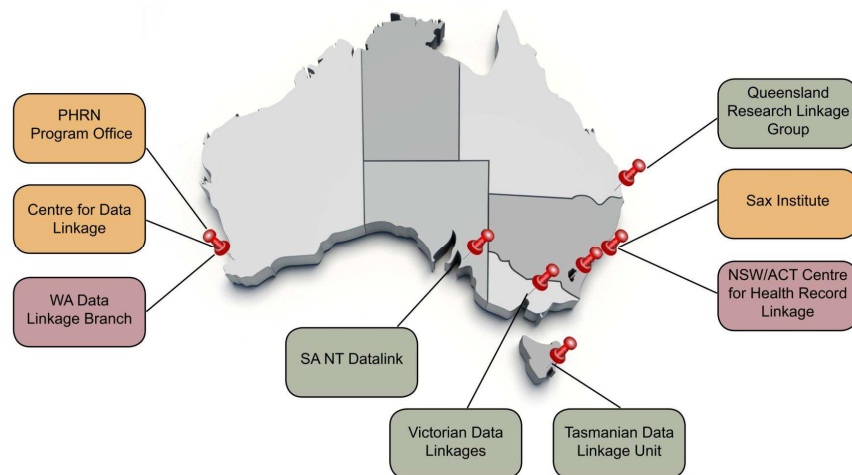


Who We Are?

Centre for Victorian Data Linkage (CVDL)



Population Health Research Network data linkage locations



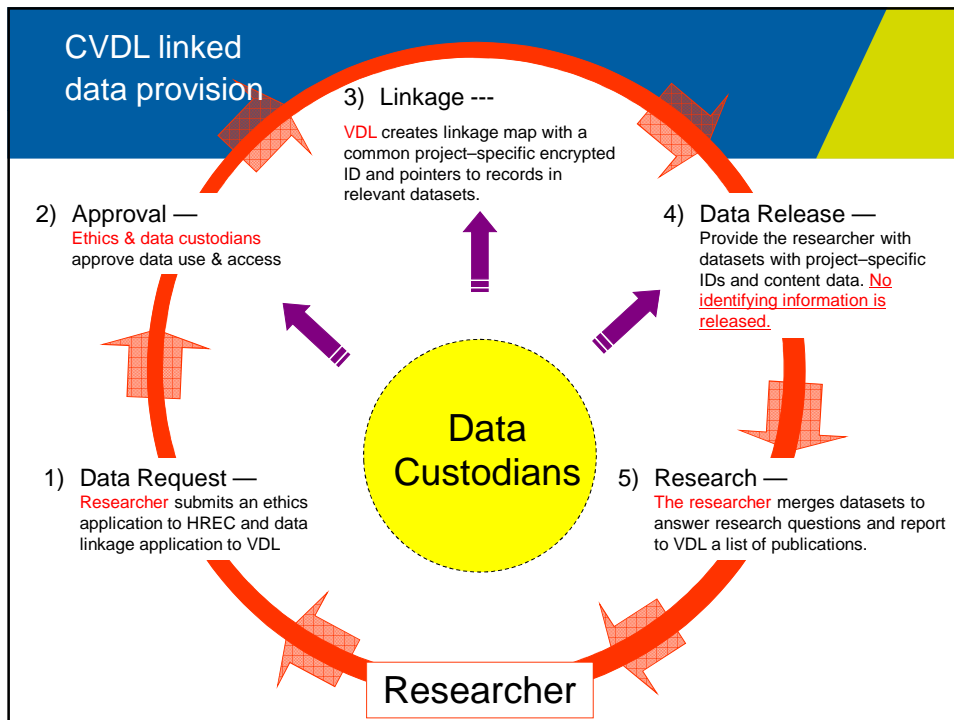
Who is involved in linkage in Victoria?

Centre for Victorian Data Linkage – creates linkage IDs which allow data to be linked within and between data collections and are involved in the governance of the project.

Data custodians – stewards for data collections, responsible for collection, use and disclosure.

Researchers/policy developers – use the linked data for analysis and research after going through an appropriate application and approval process

Ethics Committees – External researchers must gain ethics committee approval for linkage projects



What is data linkage?

- Data linkage compares records across/within datasets to identify records believed to belong to the same individual
- Enables:
 - information and evidence held within a dataset or in disparate datasets to be integrated
 - new and important questions to be researched and answered by preserving facts but protecting identity
 - patterns, trends, pathways and profiles to be identified
 - better use of administrative data as a resource for research, policy development, service and program planning

What does CVDL do?

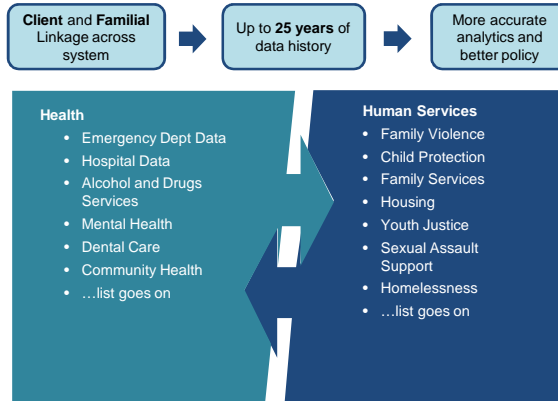
CVDL since its establishment has focused on

- Undertaking linkage projects for internal and external stakeholders
- Testing linkage concepts and developing data linkage methods
- Development of linkage infrastructure, such as the Victorian Linkage Map and the Better Patient Data database
- Developing privacy policies, protocols and procedures

Legislative framework for data linkage

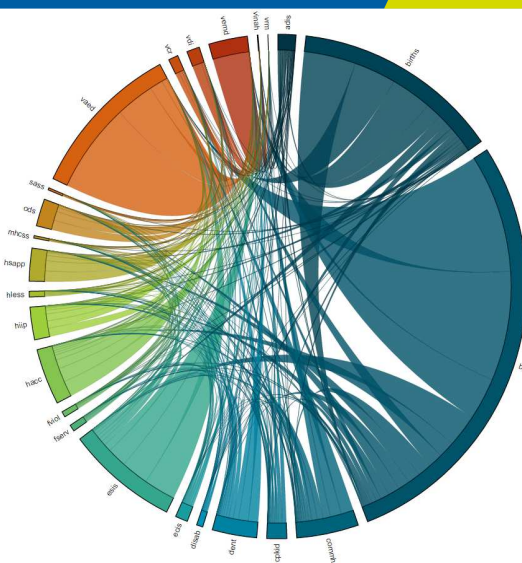
- **The legal framework is complex**
- **Disclosure of personal identifiers to the data linkage unit attracts privacy protecting principles and provisions of :**
 - Authorising legislation (eg Health Services Act 1988)
 - Privacy legislation
 - Confidentiality or data use agreements
- **Legality of linkage of any two data collections generally determined on a case by case basis**
- **Particularly complex between sectors and jurisdictions**

The Victorian Linkage Map



Victorian Data Linkage Map (VLM)

- System of links created by combining personal identifiers
- Clusters of records from 27 different datasets
- Over 100 million records



Datasets in the VLM

Better Patient Data Medicare and patient tables

Victorian Admitted Episodes Data Set
 Victorian Emergency Minimum Dataset
 Public mental health services
 Alcohol and Drug Information System
 Victorian Integrated Non-Admitted Health Dataset
 Elective Surgery Information System
 Victorian Cancer Registry Victorian Radiotherapy Minimum Dataset
 Mental Health Community Support Services
 Family Services
 Family Violence
 Sexual Assault
 Disability Services

Youth Justice

Homelessness
 Victorian Death Index
 Community Health
 Child Protection
 Public Housing Tenancies
 Home and Community Care
 Dental Health Program Dataset
 Early childhood intervention
 Births Registry
 Public housing applications
 Cradle to Kinder program
 Early Parenthood Centers

Data Linkage Concept – Hospital data

Rec_ID	First Name	Surname	DoB	Gender	Diagnosis
3	Bob	Miller	16/8/1965	Male	C460
5	Jane	Smith	18/9/1971	Female	Z088
9	Terry	Soloff	19/1/1945	Male	K920

Rec_ID	First Name	Surname	DoB	Gender
3	Bob	Miller	16/8/1965	Male
5	Jane	Smith	18/9/1971	Female
9	Terry	Soloff	19/1/1945	Male

Rec_ID	Diagnosis
3	C460
5	Z088
9	K920

← To Linkage Map

To Content Integration →

Data Linkage Concept – Death Data

ID	First Name	Surname	DoB	Gender	Date of Death
X79	Robert	Miller	16/8/1965	Male	17/9/2016
X85	Jane	Smith	18/9/1971	Female	18/5/2001
X1112	Kirsten	Bell	02/02/1993	Female	3/2/2015

ID	First Name	Surname	DoB	Gender	Rec_ID	Date of Death
X79	Robert	Miller	16/8/1965	Male	X79	17/9/2016
X85	Jane	Smith	18/9/1971	Female	X85	18/5/2001
X1112	Terry	Soloff	19/1/1945	Male	X1112	3/2/2015

← To Linkage Map

To Content Integration →

Data Linkage Concept – Linkage Map

Rec_ID	First Name	Surname	DoB	Gender
3	Bob	Miller	16/8/1965	Male
5	Jane	Smith	18/9/1971	Female
9	Terry	Soloff	19/1/1945	Male

Dataset 1

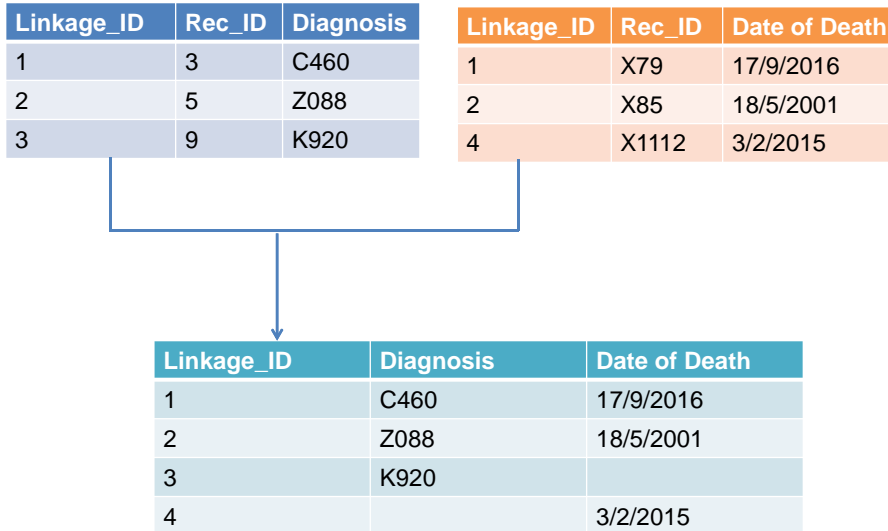
ID	First Name	Surname	DoB	Gender
X79	Robert	Miller	16/08/1965	Male
X85	Jane	Smith	18/09/1971	Female
X1112	Kirsten	Bell	02/02/2001	Female

Dataset 2

Linkage_ID	Rec_ID	ID
1	3	X79
2	5	X85
3	9	
4		X1112

Linkage Map

Data Linkage Concept – Data Integration



CVDL Systems

M10PC_1 TABLE

Summary Data Fields Data Model Graph

Filter: Maximum Rows : 500 | Field Filter

Go to row: 1

	unique_id	pcode	dobc	daybirth	mobirth	yearbirth	sex	forename	surname	fullname
1	100000006	6071	19991720	20	17	1999	1	MARK	WIGHT	MARK WIGHT
2	100000012	6031	19500226	26	02	1950	1	TIMOTHY	NOCCIOLINO	TIMOTHY NOCCIOLINO
3	100000016	6018	19570815	15	08	1957	1	VINKO	CLARKE	VINKO CLARKE
4	100000021	6530	19901108	08	11	1990	1	RODNEY	HORNSBY	RODNEY HORNSBY
5	100000029	6021	19400708	08	07	1940	1	CLINTON	THOCMAS	CLINTON THOCMAS
6	100000034	6066	19840511	11	05	1984	1	SIMON	BURROW	SIMON BURROW
7	100000038	6157	19291231	31	12	1929	1	RODNEY	CHAN	RODNEY CHAN
8	100000039	6239	19821003	03	10	1982	1	ADRIAN	HARRT	ADRIAN HARRT
9	100000041	6720	19540930	30	09	1954	1	DEREK	ABELHA	DEREK ABELHA
10	100000057	6061	19640610	10	06	1964	1		MOORE	
11	100000094	6168	19830927	27	09	1983	1	SHANE	SHARP	SHANE SHARP
12	100000099	6485	19340291	91	02	1934	1	SHANE		
13	100000109	6065	19511104	04	11	1951	1	RYAN	GREGAN	RYAN GREGAN
14	100000113	6056	19670416	16	04	1967		JAMEZ	DANIELS	JAMEZ DANIELS
15	100000116	6566	19840226	26	02	1984	1	ANTONY	LOPRESTI	ANTONY LOPRESTI
16	100000133	6311	19950609	09	06	1995	1	LESLIE	FISHER	LESLIE FISHER
17	100000139	6014	19180913	13	09	1918	1	JUSTIN	WILSON	JUSTIN WILSON
18	100000150	6065	19840417	17	04	1984	1	IAN	THOMPSON	IAN THOMPSON
19	100000153	6173	19310718	18	07	1931	1	HANK	OMOND	HANK OMOND
20	100000164	6164	19620427	27	04	1962	1	TROY	CHRISTIANO	TROY CHRISTIANO
21	100000180	6152	19960805	05	08	1996	1	ROBBIE	COOPER	ROBBIE COOPER
22	100000181	6525	19511119	19	11	1951	1	PAU	LOVLEL	PAU LOVLEL

CVDL Systems

M10PC_2

Filter: Maximum Rows: 500

UNIQUE_ID	PCODE	DOBC	DAYBIRTH	MOBIRTH	YEARBIRTH	SEX	FORENAME	SURNAME	FULLNAME	FULLNAME_MATCHCODE	CLUSTER_ID1
19	1000071282	6056	19670416	16	04	1967	JAMES DANIELS	DANIELS	JAMES DANIELS	3&B&@W\$SSSSSSSSC&@_4SSSSSS	13
20	1000051345	6056	19670416	16	04	1967	JAMES DANIELS	DANIELS	JAMES DANIELS	3&B&@W\$SSSSSSSSC&@_4SSSSSS	13
21	1000000113	6056	19670416	16	04	1967	JAMEZ DANIELS	DANIELS	JAMEZ DANIELS	3&B&@W\$SSSSSSSSC&@_4SSSSSS	13
22	1000084796	6566	19840426	26	04	1984	ANTONY LOPRESTI	LOPRESTI	ANTONY LOPRESTI	W&M\$4-7SSSSSS-@PRSSSSSSSS	14
23	1000000116	6566	19840226	26	02	1984	ANTONY LOPRESTI	LOPRESTI	ANTONY LOPRESTI	W&M\$4-7SSSSSS-@PRSSSSSSSS	14
24	1000129496	6311	19950609	09	06	1995	LESLIE FISHER	FISHER	LESLIE FISHER	G&42&Y\$SSSSSSSSW&4W7SSSSSS	15
25	1000000133	6311	19950609	09	06	1995	LESLIE FISHER	FISHER	LESLIE FISHER	G&42&Y\$SSSSSSSSW&4W7SSSSSS	15
26	1000000139	6014	19180913	13	09	1918	JUSTIN WILSON	WILSON	JUSTIN WILSON	L&W4&B\$SSSSSSSSC#4-SSSSSSSS	16
27	1000144845	6014	19180913	13	09	1918	JUSTIN WILSON	WILSON	JUSTIN WILSON	L&W4&B\$SSSSSSSSC#4-SSSSSSSS	16
28	1000169395	6014	19180913	13	09	1918	JUSTIN WILSON	WILSON	JUSTIN WILSON	L&W4&B\$SSSSSSSSC#4-SSSSSSSS	16
29	1000123961	6065	19840417	17	04	1984	IAN THOMPSON	THOMPSON	IAN THOMPSON	-2&B&A&B\$SSSSSSSS&B\$SSSSSSSS	17
30	1000142250	6065	19840417	17	04	1984	IAN THOMPSON	THOMPSON	IAN THOMPSON	-2&B&A&B\$SSSSSSSS&B\$SSSSSSSS	17
31	1000000150	6065	19840417	17	04	1984	IAN THOMPSON	THOMPSON	IAN THOMPSON	-2&B&A&B\$SSSSSSSS&B\$SSSSSSSS	17
32	1000000153	6173	19310718	18	07	1931	HANK OMOND	OMOND	HANK OMOND	#B&B\$SSSSSSSSSS2P\$RSSSSSSSS	18
33	1000152595	6173	19310718	18	07	1931	HENRY OMOND	OMOND	HENRY OMOND	#B&B\$SSSSSSSSSS2P\$RSSSSSSSS	18
34	1000000164	6164	19620427	27	04	1962	TROY CHRISTIANO	CHRISTIANO	TROY CHRISTIANO	3Y&4-@B@5SSSS-@RSSSSSSSS	19
35	1000087284	6152	19960805	05	08	1996	ROBERT COOPER	COOPER	ROBERT COOPER	3&M&Y\$SSSSSSSSM@M\$SSSSSSSS	20
36	1000000180	6152	19960805	05	08	1996	ROBBIE COOPER	COOPER	ROBBIE COOPER	3&M&Y\$SSSSSSSSM@M\$SSSSSSSS	20
37	1000051549	6152	19960805	05	08	1996	ROBERT COOPER	COOPER	ROBERT COOPER	3&M&Y\$SSSSSSSSM@M\$SSSSSSSS	20
38	1000000181	6525	19511119	19	11	1951	PAU LOVLEL	LOVLEL	PAU LOVLEL	W&M&W\$SSSSSSSSS\$N\$SSSSSSSS	21
39	1000041078	6060	19200507	07	05	1920	SCOTT THOMPSON	THOMPSON	SCOTT THOMPSON	-2&B&A&B\$SSSSSSSS4J-SSSSSSSS	22
40	1000172910	6060	19200507	07	05	1920	SCOTT THOMPSON	THOMPSON	SCOTT THOMPSON	-2&B&A&B\$SSSSSSSS4J-SSSSSSSS	22
41	1000000190	6060	19200507	07	05	1920	SCOTT THOMPSON	THOMPSON	SCOTT THOMPSON	-2&B&A&B\$SSSSSSSS4J-SSSSSSSS	22
42	1000069775	6169	19940422	22	04	1994	MATTHEW GREGAN	GREGAN	MATTHEW GREGAN	FY&F&B\$SSSSSSSSB&-SSSSSSSS	23
43	1000000197	6169	19940422	22	04	1994	MATTHEW GREGAN	GREGAN	MATTHEW GREGAN	FY&F&B\$SSSSSSSSB&-SSSSSSSS	23
44	1000006124	6169	19940422	22	04	1994	MAT GREGAN	GREGAN	MAT GREGAN	FY&F&B\$SSSSSSSSB&-SSSSSSSS	23
45	1000159161	6169	19940422	22	04	1994	MATTHEW GREGAN	GREGAN	MATTHEW GREGAN	FY&F&B\$SSSSSSSSB&-SSSSSSSS	23

CVDL Systems

M10PC_2

Filter: Maximum Rows: 500 | Row filter: CLUSTER_ID1

UNIQUE_ID	PCODE	DOBC	DAYBIRTH	MOBIRTH	YEARBIRTH	SEX	FORENAME	SURNAME	FULLNAME	FULLNAME_MATCHCODE	CLUSTER_ID1
1	1000086701	6155	19330416	16	04	1933	SIMON TANG	TANG	SIMON TANG	-8&F\$SSSSSSSSSS4&B&B\$SSSSSSSS	2852
2	1000101857	6155	19330417	17	04	1933	SIMON TANG	TANG	SIMON TANG	-8&F\$SSSSSSSSSS4&B&B\$SSSSSSSS	2852
3	1000030307	6155	19330416	16	04	1933	SIMON TANG	TANG	SIMON TANG	-8&F\$SSSSSSSSSS4&B&B\$SSSSSSSS	2852
4	1000100587	6149	19990619	19	06	1999	KIM POTTER	POTTER	KIM POTTER	N&-8Y\$SSSSSSSSSS37B\$SSSSSSSS	2853
5	1000076726	6149	19990619	19	06	1999	KIM POTTER	POTTER	KIM POTTER	N&-8Y\$SSSSSSSSSS37B\$SSSSSSSS	2853
6	1000030316	6149	19990619	19	06	1999	KIM POTTER	POTTER	KIM POTTER	N&-8Y\$SSSSSSSSSS37B\$SSSSSSSS	2853
7	1000156369	6149	19990619	19	06	1999	KIM POTTER	POTTER	KIM POTTER	N&-8Y\$SSSSSSSSSS37B\$SSSSSSSS	2853
8	1000039148	6008	19220713	13	07	1922	KEITH LEFROY	LEFROY	KEITH LEFROY	W&M\$Y&7\$SSSSSSSS3&-2SSSSSSSS	2854
9	1000030317	6020	19220713	13	07	1922	KEITH LEFROY	LEFROY	KEITH LEFROY	W&M\$Y&7\$SSSSSSSS3&-2SSSSSSSS	2854
10	1000030338	6324	19970325	25	03	1997	DANIEL BURLLES	BURLLES	DANIEL BURLLES	M&Y&W&4\$SSSSSSSSS&P\$W\$SSSSSSSS	2855
11	1000030344	6172	19970325	25	03	1997	ALFREDO WAKELY	WAKELY	ALFREDO WAKELY	L&3&W7\$SSSSSSSSS&W\$SSSSSSSSSS	2856
12	1000030348	6157	19890201	01	02	1989	WEI GLEGHORN	GLEGHORN	WEI GLEGHORN	FW&F&Y\$SSSSSSSSS&L&7\$SSSSSSSS	2857
13	1000180107	6157	19890201	01	02	1989	WEI GLEGHORN	GLEGHORN	WEI GLEGHORN	FW&F&Y\$SSSSSSSSS&L&7\$SSSSSSSS	2857
14	1000030352	6019	19591227	27	12	1959	SHAHE WATSLN	WATSLN	SHAHE WATSLN	L&-4W&B\$SSSSSSSS4&2&2SSSSSSSS	2858
15	1000030366	6018	19591227	27	12	1959	ROBERT CRANE	CRANE	ROBERT CRANE	3Y&B\$SSSSSSSSSSM@M\$SSSSSSSS	2859
16	1000030375	6330	19250924	24	09	1925	ANDREW SMAILES	SMAILES	ANDREW SMAILES	4B&W&4\$SSSSSSSSSSG&4&V&B\$SSSS	2860
17	1000030382	6252	19760216	16	02	1976	FRESERICK LUKE	JAEGER	FRESERICK LUKE	C&B&F&Y\$SSSSSSSSSSW#3\$SSSSSSSS	2861
18	1000030390	6233	19960427	27	04	1996	LUKE JAEGER	JAEGER	LUKE JAEGER	C&B&F&Y\$SSSSSSSSSSW#3\$SSSSSSSS	2862
19	1000102547	6233	19960427	27	04	1996	LUKE JAEGER	JAEGER	LUKE JAEGER	C&B&F&Y\$SSSSSSSSSSW#3\$SSSSSSSS	2862
20	1000085129	6233	19960427	27	04	1996	LUKE JAEGER	JAEGER	LUKE JAEGER	C&B&F&Y\$SSSSSSSSSSW#3\$SSSSSSSS	2862
21	1000030406	6011	19920507	07	05	1992	MICHAEL KEYS	KEYS	MICHAEL KEYS	3&4\$SSSSSSSSSSS&B73\$SSSSSSSS	2863
22	1000030421	6076	19570409	09	04	1957	WILLDAM WELCH	WELCH	WILLDAM WELCH	L&WIZ\$SSSSSSSSSS&L&W&B&B\$SSSSSS	2864
23	1000127180	6172	19280926	26	09	1928	SAMUEL PARTINGTON	PARTINGTON	SAMUEL PARTINGTON	N&Y-&B&F--&B\$SSSSSS4&B\$SSSSSSSS	2865
24	1000030445	6172	19280926	26	09	1928	SAMUEL PARTINGTON	PARTINGTON	SAMUEL PARTINGTON	N&Y-&B&F--&B\$SSSSSS4&B\$SSSSSSSS	2865
25	1000049193	6172	19280926	26	09	1928	SAMUEL PARTINGTON	PARTINGTON	SAMUEL PARTINGTON	N&Y-&B&F--&B\$SSSSSS4&B\$SSSSSSSS	2865

Examples of publications based on linked data

- Adverse outcomes associated with elective knee arthroscopy: a population-based cohort study
- Using linked data to more comprehensively measure the quality of care for stroke
- Clinical presentation and patterns of care for short-term survivors of malignant glioma
- Myelodysplastic syndrome incidence, transfusion dependence, health care use, and complications: an Australian population-based study 1998 to 2008
- Palliative care for patients with chronic obstructive pulmonary disease: Exploring the landscape.
- Emergency department access targets and the older patient: a retrospective cohort study of emergency department presentations by people living in residential aged care facilities.
- Factors involved in intensive care unit mortality following medical retrieval: identifying differences between intensive care unit survivors and non-survivors.

Projects

- "Air Pollution Exposure and Health Outcomes: Evidence from the Changes in Energy Infrastructure and Environmental Policy in Victoria"
- Understanding the health service use patterns of children in Out-of-Home-Care compared to local and State averages of all children
- Transitioning from hospital to home: Effectiveness of occupational therapy home visit discharge planning program for at risk older adults
- Outcomes following prostate cancer treatment
- Family Violence Vic Pol and CMIODS
 - Examining the relationship between family violence and mental health
- HealthLinks Chronic Care
 - Ongoing identification of hospital inpatients who are at risk of multiple admissions over the following 12 months to support a new funding model
- Burns and Psychiatry comorbidity

To find out more...

www.dhhs.vic.gov.au/vdl

VDL Helpdesk: vdI@dhhs.vic.gov.au

Tel: 9096 3018