Statistical methods for multi-jurisdictional Australian vaccine safety investigations of rare adverse events

Hannah J Morgan[+1,2,3], Lauren Bloomfield[+4,5], Hazel J Clothier[1,2,3], Sera Ngeh[4], Gemma Cadby[4], Dale Carcione[4], James H Boyd[6], Gonzalo Sepulveda Kattan[2,3], Jim P Buttery*[1,2,3,7], Paul Effler*[4]

[1] The Department of Paediatrics, The University of Melbourne, Melbourne, Australia.

[2] Epi-Informatics, Centre for Health Analytics, Melbourne Children's Campus, Melbourne, Australia.

[3] Epi-Informatics Group and SAEFVIC, Murdoch Children's Research Institute, Melbourne, Australia.

[4] Department of Health, Western Australia, Australia.

[5] The University of Notre Dame, Fremantle, Western Australia, Australia

[6] Department of Public Health, La Trobe University, Melbourne, Victoria, Australia.

[7] Infectious Diseases Unit, Department of General Medicine, Royal Children's Hospital Melbourne, Melbourne, Victoria, Australia.

[+] Joint first authors

* Joint supervising authors

Corresponding author:

Hannah J Morgan

Email: hannah.morgan@mcri.edu.au

Name	**ORCID ID**
Hannah Morgan	0000-0001-8717-908X
Lauren Bloomfield	0000-0001-5890-4254
Hazel Clothier	0000-0001-7594-0361
Sera Ngeh	0000-0001-9565-7846
Gemma Cadby	0000-0001-7317-6531
Dale Carcione	0009-0008-3655-7587
James Boyd	0000-0002-8925-8811
Gonzalo Sepulveda Kattan	0009-0002-4960-326X
Jim Buttery	0000-0001-9905-2035
Paul Effler	0000-0002-8890-4201

Keywords

Vaccine safety, Pharmacoepidemiology, data linkage, immunisation

Acknowledgements

The authors would like to acknowledge the Victorian Department of Health as the source of VPDC and VAED data for this study, and the Centre for Victorian Data Linkage for the provision of data linkage. The authors would also like to thank the staff at the Western Australian Data Linkage Branch, and Data Custodians of the Hospital Morbidity Data Collection and Australian Immunisation Register for providing linked health data.

Declarations

Funding – SAEFVIC and its Vaccine Safety Health Link program were funded in whole for the period of this work by the Department of Health, Victoria. WAVSS is funded in whole by the Department of Health, Western Australia.

Conflicts of interest – The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this study.

Availability of data and material – All data used in this study is sensitive and are not available for distribution.

Ethics approval – All work conducted in VSHL is overseen by The Royal Children's Hospital Melbourne Ethics Committee (HREC 79964). All vaccine safety work conducted using the WA VLDR is conducted with approval from the relevant Data Custodians for the purposes of public health surveillance.

Consent to participate - The VSHL platform uses whole-of-population routinely collected and de-identified data. Therefore, a waiver of consent was sought and approved for the VSHL project in line with Australia's National Statement section 2.3.10.

Consent for publication – NA

Code availability – R code is available upon request

Author contributions – All authors have contributed to the development of this article.,. All authors have read and approved the final version, and take responsibility for its contents.

- Concept and design: All authors
- Statistical analysis: HM, LB, GK, GC
- Drafting of the manuscript: HM, LB
- Critical revision of the manuscript for important intellectual content: All authors
- Administrative, technical, or material support: HM, LB
- Supervision: JPB, JHB, HJC

Table 1: Linked database comparison between Victoria and Western Australia.

Both sites have linked repositories of immunisation and health outcome data – the Vaccine Safety Health Link (VSHL) for Vic. and the Vaccine Linked Data Repository (VLDR) for WA. These records contain rich demographic and medical information to provide a rapid and sensitive investigation of vaccine safety concerns with less reporting bias.

	Victoria (VSHL)	**Western Australia (VLDR)**
Organisation	SAEFVIC which is funded by the Department of Health, Victoria.	The Department of Health, Western Australia.
Population size	6.68 million as of 2020	2.67 million as at 2020
Datasets		
Immunisation	Australian Immunisation Register (AIR)	Australian immunisation register (AIR)
Hospital - Emergency	Victorian Emergency Minimum Dataset (VEMD)	Emergency Department Data Collection (EDDC)
Hospital - Admissions	Victorian Admitted Episodes dataset	Hospital Morbidity Data Collection (HMDC)
Mortality	Victorian Deaths Index (VDI)	WA Mortality Register
Notifiable disease	Public Health Event Surveillance System (PHESS) for non-COVID-19 diseases and The Transmission Response Epidemiology Victoria Information System (TREVI) for COVID-19 disease.	No notifiable disease dataset was used for vaccine safety.
Pregnancy and Birth	Victorian Perinatal Data Collection (VPDC)	Midwives Notification System (MNS)
Data format		
Identifiers	Anonymous unique identifiers	Personal details (e.g. name, date of birth)
Refresh frequency	The individual datasets contributing to VSHL update as frequently as daily (AIR, VEMD) and as infrequently as 18-monthly (VPDC). The VSHL dataset itself is re-linked every month with whichever individual datasets have been updated since the previous refresh.	The individual datasets contributing to VLDR update as frequently as daily (AIR) and as infrequently as 12-monthly (MNS). The VLDR dataset is linked in near real-time as the data becomes available.

Table 2: Guillain-Barré syndrome alignment to combined analysis criteria

Number	Criteria	Guillain-Barré syndrome (GBS) alignment
1	The condition is rare, or the number of doses administered for the vaccine in question is too low to have generated sufficient case numbers in single jurisdictions alone	GBS is a very rare condition with a background rate of 2-3 cases per 100,000 per year. GBS is a known vaccine safety signal following viral vector-based vaccines, however, WA have not been able to detect the association due to insufficient ChAdOx1-S doses administered and small case numbers [22, 39].
2	An understanding of the diagnostic practices of the AEFI and the similarities and differences of their application between jurisdictions.	GBS has a clear ICD-10-AM diagnostic code that is applied consistently throughout Australia.
3	A power analysis determines that between the participating jurisdictions there is sufficient sample size to statistically measure the association.	The *samplesize* function in the SCCS package was used to determine minimum sample size for $\alpha = 0.05$ and $1-\beta = 0.8$. A total of 402 cases were needed to detect RI ≥ 2, which was feasible with the combined case numbers in each jurisdiction, but not by WA alone.
4	The vaccine safety question is answerable with the data fields available.	The relationship between COVID-19 vaccination exposure and GBS outcome is able to be answered using a self-controlled case series methodology that uses the following available data fields: date of birth, COVID-19 vaccination administered, vaccination date, GBS ICD-10-AM code in hospital, GBS ICD-10-AM hospitalisation date, state.
5	The vaccine safety question is of mutual interest between the vaccine safety jurisdictions and the communities they serve.	GBS is an adverse event of special interest and has been investigated by the Therapeutic Goods Administration since May 2021. It is an AEFI of national interest that has impacted Vic. and WA populations [40].
6	The vaccine safety question falls within the existing public health surveillance agreements that each jurisdiction has in place or there is access to timely approval processes. Jurisdictions should aim to have pre-approval to conduct analysis for any emerging research question of public health interest ahead of time to ensure timely response to public health questions.	Vic. and WA have existing approvals to conduct investigations into questions of public health importance.

Table 3: Self-controlled case series results including relative incidence (RI), 95% confidence intervals (95%CI) and p-values (P) for Victoria and Western Australia separately, and combined analysis results calculated using strategy 1 and strategy 2. Results are disaggregated by COVID-19 vaccine type, vaccine dose, sex and 50-year age group.

	Victoria			**Western Australia**				**Meta-analysis (strategy 1)**		**Combined linelist (strategy 2)**	
	Count included in analysis	**RI(95%CI), P**	**SE**	**Count included in analysis**	**RI(95%CI), P**	**SE**	**Combined count included in analysis**	**RI(95%CI), P**	**SE**	**RI(95%CI), P**	**SE**
Vaccine factors											
All COVID-19 vaccines	519	**1.50 (1.16, 1.95) p0.002**	0.132	176	1.03 (0.62,1.73) p0.91	0.264	695	**1.39 (1.10,1.75) p0.005**	0.118	**1.32 (1.04, 1.68) p 0.02**	0.121
mRNA	473	0.89 (0.61, 1.30) p0.55	0.193	171	1.11 (0.61, 1.99) p0.74	0.301	644	0.95 (0.69,1.31) p0.75	0.163	0.89 (0.64, 1.23) p0.48	0.168
Dose 1	246	1.14 (0.56, 2.30) p0.72	0.360	100	1.05 (0.33,3.30) p0.94	0.586	346	1.11 (0.61,2.03) p0.72	0.307	0.91 (0.47, 1.76) p0.77	0.338
Dose 2	266	1.05 (0.52, 2.12) p0.90	0.360	104	1.00 (0.32,3.16) p0.99	0.586	370	1.04 (0.57,1.89) p0.91	0.307	0.95 (0.51, 1.78) p0.87	0.321
Dose 3	367	0.66 (0.31, 1.49) p0.43	0.339	142	0.73 (0.23,2.29) p0.59	0.584	509	0.68 (0.38,1.20) p0.18	0.293	0.69 (0.37, 1.29) p0.24	0.319
Dose 4	191	0.72 (0.26, 1.95) p0.52	0.505	64	1.66 (0.52,5.30) p0.39	0.591	255	1.02 (0.48,2.17) p0.95	0.384	0.82 (0.37, 1.85) p0.64	0.413
Viral vector	268	**2.95 (2.09, 4.16) p<0.0001**	0.176	72	1.02 (0.37, 2.79) p0.98	0.515	340	**2.64 (1.90,3.66) p<0.0001**	0.167	**2.45 (1.76, 3.41) <0.0001**	**0.169**
Dose 1	266	**4.44 (3.10, 6.40) p<0.0001**	0.181	72	0.48 (0.07, 3.43) p0.46	1.007	338	**4.14 (2.92,5.87) p <0.0001**	0.178	**4.89 (2.73, 5.55) <0.0001**	**0.181**
Dose 2	222	0.31 (0.08, 1.24) p0.09	0.710	65	1.64 (0.51, 5.21) p0.41	0.591	287	0.83 (0.34,2.02) p0.68	0.454	0.62 (0.25, 1.49) p0.28	0.451
Dose 3	13	2.81 (0.36, 21.65) p0.32	1.041	0	Not enough data	NA	13	NA	NA	3.75 (0.48, 29.63) p0.21	1.054
Dose 4	3	Not enough data	NA	0	Not enough data	NA	3	Not enough data	NA	Not enough data	NA
Protein subunit	10	Not enough data	NA	11	Not enough data	NA	21	Not enough data	NA	Not enough data	NA
Demographic factors											
Sex											

Male	295	1.22 (0.84, 1.78) p0.30	0.197	98	1.50 (0.82,2.76) p0.19	0.310	393	1.29 (0.93,1.79) p0.12	0.166	1.27 (0.92, 1.77) p0.14	0.165
Female	224	**1.88 (1.31, 2.70) p<0.0001**	0.184	78	0.53 (0.19,1.45) p0.22	0.515	302	**1.63 (1.16,2.29) p0.0049**	0.173	**1.51 (1.06, 2.13) p0.02**	0.178
Age group											
<=50	190	1.05 (0.61, 1.82) p0.85	0.279	66	1.57 (0.72,3.45) p0.26	0.401	256	1.20 (0.76,1.88) p0.43)	0.229	1.15 (0.73, 1.82) p0.55	0.234
> 50	329	**1.71 (1.27, 2.30) p<0.0001**	0.151	111	0.81 (0.41,1.60) p0.54	0.349	440	**1.52 (1.16,1.99) p0.0025**	0.139	**1.48 (1.12, 1.95) p0.006**	0.142

Table 4: Sensitivity analysis. Jurisdiction-specific self-controlled case series results including relative incidence (RI), 95% confidence intervals (95%CI) and p-values (P) – using dose sequence according to COVID-19 vaccines administered and using dose sequence according to platform administered.

Victoria

	Dose sequenced along all vaccines (As per Table 3)			Dose sequenced along that specific vaccine type.		
	Count	**RI(95%CI) p-value**	**se**	**Count**	**RI(95%CI) p-value**	**se**
All COVID-19 vaccines	519	1.50 (1.16, 1.95) p0.002	0.132	519	1.50 (1.16, 1.95) p0.002	0.132
mRNA	473	0.89 (0.61, 1.30) p0.55	0.193	473	0.89 (0.61, 1.30) p0.55	0.193
Dose 1	246	1.14 (0.56, 2.30) p0.72	0.360	473	0.97 (0.46,2.06) p0.94	0.383
Dose 2	266	1.05 (0.52, 2.12) p0.90	0.360	268	0.91 (0.43,1.92) p0.78	0.383
Dose 3	367	0.66 (0.31, 1.49) p0.43	0.339	360	0.67 (0.32,1.42) p0.30	0.382
Dose 4	191	0.72 (0.26, 1.95) p0.52	0.505	184	0.56(0.18,1.75) p0.319	0.582
Viral vector	268	2.95 (2.09, 4.16) p<0.0001	0.176	268	2.95 (2.09, 4.16) p<0.0001	0.176
Dose 1	266	4.44 (3.10, 6.40) p<0.0001	0.181	268	4.44 (3.10, 6.40) p<0.0001	0.181
Dose 2	222	0.31 (0.08, 1.24) p0.09	0.710	214	0.31 (0.08,1.28) p0.11	0.710
Dose 3	13	2.81 (0.36, 21.65) p0.32	1.041	10	3.75 (0.48,29.63) p0.21	1.054
Dose 4	3	Not enough data	NA			
Protein subunit	10	Not enough data	NA	10	Not enough data	NA

Western Australia

	Dose sequenced along all vaccines (As per Table 3)			Dose sequenced along that specific vaccine type.		
	Count	**RI(95%CI) p-value**	**se**	**Count**	**RI(95%CI) p-value**	**se**
All COVID-19 vaccines	176	1.03 (0.62,1.73) p0.91	0.264	176	1.03 (0.62,1.73) p0.91	0.264
mRNA	171	1.11 (0.61, 1.99) p0.74	0.301	171	1.11 (0.61, 1.99) p0.74	0.301
Dose 1	100	1.05 (0.33,3.30) p0.94	0.586	171	1.02 (0.42,2.48) p0.97	0.454
Dose 2	104	1.00 (0.32,3.16) p0.99	0.586	151	1.40(0.62,3.16) p0.42	0.417
Dose 3	142	0.73 (0.23,2.29) p0.59	0.584	80	0.43 (0.06,3.10) p0.40	1.006
Dose 4	64	1.66 (0.52,5.30) p0.39	0.591	13	Not enough data	NA
Viral vector	72	1.02 (0.37, 2.79) p0.98	0.515	72	1.02 (0.37, 2.79) p0.98	0.515
Dose 1	72	0.48 (0.07,3.43) p0.46	0.476	72	0.48 (0.07,3.43) p0.46	0.476
Dose 2	64	1.64 (0.51,5.21) p0.41	0.591	65	1.64 (0.51,5.21) p0.41	0.591

Dose 3	0	Not enough data	NA	0	Not enough data	NA
Dose 4	0	Not enough data	NA	0	Not enough data	NA
Protein subunit	11	Not enough data	NA	11	Not enough data	NA

Table 5: Sensitivity analysis. Self-controlled case series fixed- and random-effects meta-analysis (strategy 1) results including relative incidence (RI), 95% confidence intervals (95%CI) and p-values (P). Results are disaggregated by COVID-19 vaccine type, vaccine dose, sex and 50-year age group.

		Meta-analysis (fixed effects)		**Meta-analysis (random effects)**	
	Combined count included in analysis	**RI(95%CI), P**	**SE**	**RI(95%CI), P**	**SE**
Vaccine factors					
All COVID-19 vaccines	695	**1.39 (1.10,1.75) p0.005**	0.118	1.33 (0.95,1.88) p0.10	0.175
mRNA	644	0.95 (0.69,1.31) p0.75	0.163	0.95 (0.69,1.31) p0.75	0.163
Dose 1	346	1.11 (0.61,2.03) p0.72	0.307	1.11 (0.61,2.03) p0.72	0.307
Dose 2	370	1.04 (0.57,1.89) p0.91	0.307	1.04 (0.57,1.89) p0.91	0.307
Dose 3	509	0.68 (0.38,1.20) p0.18	0.293	0.68 (0.38,1.20) p0.18	0.293
Dose 4	255	1.02 (0.48,2.17) p0.95	0.384	1.03 (0.46,2.33) p0.94	0.414
Viral vector	340	**2.64 (1.90,3.66) p<0.0001**	0.167	1.94 (0.70,5.36) p0.20	0.519
Dose 1	338	**4.14 (2.92,5.87) p <0.0001**	0.178	1.82 (0.21,15.42) p0.58	1.090
Dose 2	287	0.83 (0.34,2.02) p0.68	0.454	0.75 (0.15,3.81) p0.726	0.832
Dose 3	13	Not enough data	NA	Not enough data	NA
Dose 4	3	Not enough data	NA	Not enough data	NA
Protein subunit	21	Not enough data	NA	Not enough data	NA
Demographic factors					
Sex					
Male	393	1.29 (0.93,1.79) p0.12	0.166	1.29 (0.93,1.79) p0.120	0.166
Female	302	**1.63 (1.16,2.29) p0.0049**	0.173	1.09 (0.32,3.73) p0.89	0.626
Age group					
<=50	256	1.20 (0.76,1.88) p0.43)	0.229	1.20 (0.76,1.88) p0.43	0.229
> 50	440	**1.52 (1.16,1.99) p0.0025**	0.139	1.26 (0.61,2.59) p0.53	0.368