THE EUROPEAN SOCIETY OF MEDICINE
Medical Research Archives, Volume 12 Issue 7

RESEARCH ARTICLE

# The correlation between Australian Excess Deaths by State and Booster Vaccinations

**D.E. Allen**[1]

[1]School of Mathematics and Statistics, University of Sydney, NSW, Australia, School of Business and Law, Edith Cowan University, Joondalup, WA, Australia, and Department of Finance, Asia University, Taiwan.

OPEN ACCESS

**PUBLISHED**
31 July 2024

**CITATION**
Allen, DE. 2024. The correlation between Australian Excess Deaths by State and Booster Vaccinations. Medical Research Archives, [online] 12(7). https://doi.org/10.18103/mra.v12i7.5485



**DOI**
https://doi.org/10.18103/mra.v12i7.5485

**ISSN**
2375-1924

## ABSTRACT

The study explores the relationship by Australia State between COVID Booster Vaccinations and excess deaths. There is evidence of a very strong correlation in ordinary least squares regression analysis. Cross-validation tests support the strength of the regression relationship. The results suggest that it would be worthwhile to explore these associations in greater depth as it is an important public health issue.

## 1. Introduction

This paper features an analysis of factors associated with the recent experience of excess deaths, by individual state and territory in Australia as it relates to the official historical record of total booster vaccinations by State and Territory. It relates it to vaccine data obtained from Australian Government official health statistics. (Vaccine data obtained from https://www.health.gov.au/ourwork/covid-19-vaccines/vaccination-numbers-and-statistics, as accessed 1/1/2024). Additional information accessed includes figures for actual deaths, expected deaths, populations, total vaccinations, booster doses, recent vaccinations, and total vaccinations, as broken down by State and Territory, Simple ordinary least squares regression analysis (OLS), is used to analyse the degrees of correlation between the summary figures by state and territory with the focus being on excess deaths by State and Territory as the dependent variable, although other variables are also analysed.

Allen (2023) undertook a similar analysis across OECD countries using OECD statistics [1]. He reported that cross sectional regression analyses suggest that COVID-19 booster vaccines explain between 69 and 79 per cent of the variation in excess deaths in OECD countries as captured by excess deaths in the first week or averaged across the first three months of 2023. However, there are many difficulties in untangling the complex potential causal relationships between these variables. Cross-sectional regression analyses ignore sequential timing.

For example, in the case of potential sequential effects, it could be the case that contracting a COVID infection weakens the immune system and leads to subsequent immune system difficulties. Gao et. al (2023) analysed immune cell responses to Pfizer vaccinations [2]. They noted that: "In addition, the apparent damage of the CD8+ T cell response by viral infection is cause for concern and may leave even vaccinated individuals with a previous infection at risk for subsequent infections or other health issues."

Krug et al., (2021) have analysed the incidence of myocarditis and perocarditus related to vaccine damage in adolescents [3]. They conclude that in this particular age group careful assessment has to be made of the trade-off between protection against severe disease vis-à-vis damage from myo/periocarditus in what is essentially, an overall low-risk cohort of the general population.

Chenupatti et al., (2024), assessed the impact of Covid 19 through the lens of the gastrointestinal (GI) system [4]. They report that throughout the pandemic the presence of GI symptoms was recorded in a range from 28 per cent to 50 percent of patients. These effects were more long-lasting than the respiratory tract ones and were one of the characteristics of long-Covid. These effects on the GI system were present in both adults and children (aged < 18 years).

Pillay et al. (2022) adopted a machine learning approach to survey studies on the presentation and clinical course of more than 8000 reported cases of myocarditis and pericarditis and on and some initial reports of longer-term outcomes [5]. They consider possible mechanisms and discuss the incidence of myocarditis which they suggest is probably the highest in males aged 12-29 years and more likely with Moderna than with Pfizer mRNA vaccines. They observe that longer dosing intervals might be beneficial and that even though most cases of myocarditis are mild and self-limiting, the data on children and some severe cases is limited.

Mostert et al., (2024) analyse excess deaths using data from the 'Our World in Data Database' [6]. Excess mortality is assessed as a deviation between the reported number of deaths in a country during a certain time period and the expected number of deaths in a country for that period under normal conditions.

They report that the total number of excess deaths in 47 countries of the Western World was 3 098 456 from 1 January 2020 until 31 December 2022. Excess mortality was documented in 41 countries (87%) in 2020, 42 countries (89%) in 2021 and 43 countries (91%) in 2022. They suggest that this raises serious concerns, and that Government leaders and policymakers need to thoroughly investigate underlying causes of persistent excess mortality.

The author would agree with their statement because the reasons for excess deaths are open to debate. For example, the Society of Actuaries in Australia have a section of their website devoted to research papers produced by their Covid-19 Mortality Working Group

(See: https://www.actuaries.asn.au/microsites/pandemics-resource-centre/covid19-articles, accessed 9 January 2024). They report that there was no excess mortality in Australia in August 2023 and suggest that: the August figure was due to lower than predicted deaths from respiratory disease and dementia in the month, which was, in turn, the result of low prevalence of influenza (as the flu season was earlier than usual this year) and COVID-19. This was offset by higher than predicted deaths from COVID-19 and other causes. Furthermore, they suggest that whilst total excess mortality for the first eight months of 2023 was 6% (95% confidence interval: 4% to 8%) or +6,400 deaths – i.e., there were 6,400 more deaths than expected if the pandemic had not happened. They suggest that more than half of the YTD excess mortality is due to deaths from COVID-19 (+3,500 deaths), with another +1,000 COVID-19 related deaths, and the remaining excess of +2,000 deaths had no mention of COVID-19 on the death certificate.

A counter argument would be that either infection by COVID-19 or potential vaccine damage could potentially lead to cardiomyothapy or pericarditis. This could lead to subsequent morbidity via a heart-related event, for example, as recorded on a death certificate, which would not necessarily mention COVID19. Yamamoto, (2022) notes that "It has been hypothesized that there will be an increase in cardiovascular diseases, especially acute coronary syndromes, caused by the spike proteins in genetic vaccines" [7].

This brief review of the literature has suggested that there are a number of potential routes for vaccine-related damage to occur and that there is broad evidence of an increase in excess deaths following the implementation of vaccine strategies as a response to the Covid 19 pandemic. The purpose of the current paper is to explore the statistical linkage between excess deaths and Covid Booster vaccinations in an Australian context.

The paper is divided into five sections, the research methods are discussed in section 2, the data and descriptive statistics in section 3, the results and discussion in section 4, and a conclusion in section 5.

## 2. Methods.

The basic method of analysis is ordinary least squares regression analysis (OLS), as shown below.

$$y_i = a_i + bx_i + e_i.$$

This is accompanied by the usual assumptions about the linearity of the relationship between the dependent variable $y_i$, and the independent variable $x_i$, the absence of endogeneity, but given that bivariate regressions are involved, this is not a major issue, the existence of Gaussian distributions, homoscedasticity or constancy of the variance of the error term, and the fact that it is independent and identically distributed. The customary assumption is that sample of the general population used to produce the observations used in the analysis is randomly drawn.

Given that the number of observations used in the bivariate regressions applied in this study are limited by the number of Australian States and Territories, this issue was given considerable attention, and cross-validation techniques were applied.

Bates, et al., (2023) draw attention to the fact that cross-validation (CV) is a widely used technique to estimate prediction error, but suggest its behaviour is complex and mis-understood [8]. Rather than cross-validation estimating the prediction error for the model at hand, as fit to the training data, they argue this is not the case for the linear model fit by ordinary least squares and suggest that it estimates the average prediction error of models fit on other unseen training sets drawn from the same population. They introduce a nested CV scheme to estimate this variance more accurately.

Bates et al., (2023) suggest that cross-validation estimates CV should instead be treated as an estimator of the average prediction error across training sets [8]. The qualities of various related estimators are discussed by Zhang 1995, Zhang 1993, Hastie, Tibshirani, and Friedman 2009; and Yousef 2020 [9, 10, 11, 12]. See also Rosset and Tibshirani (2020) and Wager (2020) for a discussion about different potential estimands [13, 14].

In this paper, cross-validation techniques are applied together with basic descriptive statistics and heatmaps to explore the correlations between variables used in the bivariate regression analyses.

## 3. Data and descriptive statistics

The recent experience of excess deaths by Australian State, as published by the Australian Bureau of Statistics (19/07/2023)

(https://www.abs.gov.au/articles/measuring-australias-excess-mortality-duringcovid-19-pandemic-until-first-quarter-2023, accessed 9/1/2024), is shown in Table 1. This data for the first quarter of 2023 was used in the analysis, apart from in the case on the Northern Territory, for which I used the 2022 data.

Vaccine data for Australian States and Territories was obtained from https://www.health.gov.au/our-work/covid-19-vaccines/vaccination-numbers-andstatistics. Series were obtained for total deaths, expected deaths, excess deaths, total vaccinations, booster doses, recently vaccinated, populations, and the unvaccinated. A summary of these figures representing the 8 States and Territories is shown in Table 2. A large variation across the states is to be expected given the differences in the sizes of the populations. For example, NSW has a population of 8.3 million whilst the population of the Northern Territory is only 252,500. The greatest variation, in terms of the standard deviation, is in the total number of vaccinations by State and Territory, followed by the variation in populations.

**Table 1**: Excess Deaths by State

Excess mortality as a percentage above expected by jurisdiction, 2020-23

| | 2020 | 2021 | 2022 | 2023 |
|---|---|---|---|---|
| Australia | -3.1 | 1.4 | 10.9 | 9.1 |
| New South Wales | -4.1 | 0.1 | 10.7 | 8.6 |
| Victoria | -0.9 | 3.4 | 13.2 | 12 |
| Queensland | -4.3 | 0.8 | 10.1 | 7.8 |
| South Australia | -3.2 | 0.5 | 9.2 | 8.9 |
| Western Australia | -3.9 | 0.6 | 6.2 | 6.1 |
| Tasmania | -3.6 | 5.8 | 13.6 | 17.3 |
| Northern Territory | 1.5 | 6.8 | 10.6 | np |
| Australian Capital Territory | -4.3 | -2.8 | 12.1 | 8.9 |

a. Data is provisional and subject to change.
b. Years are based on a sum of ISO weeks derived from the weekly modelling.
There are 53 weeks in 2020. There are 52 weeks in 2021 and 2022.
Excess mortality has been estimated for the first 12 weeks of 2023.
c. Deaths in 2023 are deaths that occurred by 26 March and were registered and received by the ABS by 31 May 2023.

Source: Australian Bureau of Statistics, Measuring Australia's excess mortality during the COVID-19 pandemic until the first quarter 2023 19/07/2023

The variables used in the analysis, as shown in the correlation matrix in the lower section of Table 2, are highly correlated. The critical value for the correlation coefficient, in a two-tailed test across 8 variables, is 0.71. Thus, all the pairwise correlation coefficients are significant at a 5 per cent level or greater, except for all the correlation coefficients involving the unvaccinated.

The highest correlation is between excess deaths and booster vaccine doses with a value of 0.87. However, the correlations between total vaccinations, and recent vaccinations are 0.85 and 0.86 respectively. Unremarkably, there is a high correlation between excess deaths and total population with a value of 0.84.

Figure 1 displays a heatmap of the correlations between variables. The last column and the bottom row of the heatmap show that the unvaccinated have the lowest correlations with the other variables in the analysis. Nevertheless, there is still a correlation of 0.7 between the unvaccinated and population. However, this is probably to be expected given that that States and Territories with larger populations are likely to have more unvaccinated people. Most revealing is the insignificant correlation of 0.5 between the unvaccinated and excess deaths.

**Table 2**: Descriptive Statistics of State Series and Correlation Matrix

| Variable | Mean | Median | Maximum | Minimum | St. Deviation |
|---|---|---|---|---|---|
| Expected Deaths | 4624.1 | 3255.5 | 11540 | 550 | 4141.8 |
| Actual Deaths | 5124.1 | 3775.5 | 12529. | 599.00 | 4504.9 |
| Excess Deaths | 320.25 | 215.50 | 1051.0 | -268.00 | 466.37 |
| Total Vaccinations | 8,618,100 | 6,100,000 | 20,140,000 | 644,400 | 7,917,600 |
| Booster Doses | 521,180 | 365,150 | 1,300,000 | 18,700 | 477,860 |
| Population | 3,329,200 | 2,365,200 | 8,339,300 | 252,500 | 3,149,000 |
| Recently Vaccinated | 236,380 | 168,250 | 584,700 | 9800 | 218,010 |
| UnVaccinated | 70,588 | 46,150 | 208,200 | 9200 | 68963 |

Correlation coefficients, using the observations 1–8
5% critical value (two-tailed) = 0.7067 for n = 8

| ExcessDeaths | TOTVAC | BoostD | RECVAC | UNVAC | |
|---|---|---|---|---|---|
| 1.0000 | 0.8468 | 0.8671 | 0.8578 | 0.5071 | ExcessDeaths |
| | 1.0000 | 0.9979 | 0.9984 | 0.6230 | TOTVAC |
| | | 1.0000 | 0.9993 | 0.6055 | BoostD |
| | | | 1.0000 | 0.5992 | RECVAC |
| | | 1.0000 | | | UNVAC |

| POPD | COVCASES | |
|---|---|---|
| 0.8411 | 0.8282 | ExcessDeaths |
| 0.9978 | 0.9508 | TOTVAC |
| 0.9940 | 0.9609 | BoostD |
| 0.9950 | 0.9615 | RECVAC |
| 0.6658 | 0.5232 | UNVAC |
| 1.0000 | 0.9464 | POPD |
| | 1.0000 | COVCASES |

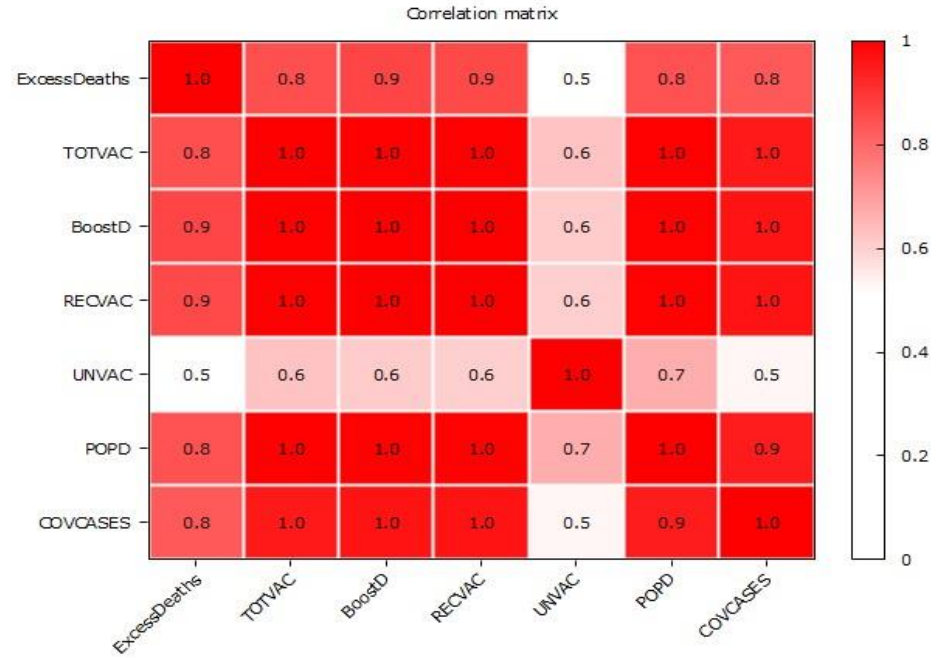



**Figure 1**. Heat Map of Correlations between Variables

## 4. Bivariate regression Results, cross-validation and Discussion

4,1 BIVARIATE REGRESSIONS

To explore the relationships between the variables, bivariate regression analyses were performed on pairs of variables using excess deaths as the dependent variable, followed by total vaccinations and booster doses. The results of these analyses are shown in Table 3, in Part A of the Table, followed by some cross-validation tests in Part B.

**Table 3**: Bivariate OLS regression Analyses and Cross-Validation Tests

Part A: Bivariate OLS regression Analyses

| **Regression** | **Coefficient** | **Adjusted R-Square** | **F-Statistic** | **F-Statistic Residuals regression** |
|---|---|---|---|---|
| Excess Deaths = $C_1 + \beta_1$BoostD | 0.000846*** | 0.71 | 18.18*** | 0.010520 |
| Excess Deaths = $C_2 + \beta_2$TotVAC | 0.0000499*** | 0.71 | 15.21*** | 0.000319 |
| Excess Deaths = $C_3 + \beta_3$RECVAC | 0.00183491*** | 0.69 | 16.71*** | 0.003507 |
| Excess Deaths = $C_4 + \beta_4$UNVAC | 0.00342938 | 0.13 | 2.08 | 3.107021 |
| Excess Deaths = $C_5 + \beta_5$COVCASES | 0.000332594** | 0.63 | 13.11** | 0.063274 |
| Excess Deaths = $C_6 + \beta_6$POP | 0.124562 *** | 0.66 | 14.51*** | 0.019782 |
| TotVAC = $C_7 + \beta_7$COVCASES | 6.38561*** | 0.98 | 321.9*** | 0.666778 |
| BoostD = $C_8 + \beta_8$COVCASES | 0.385757 *** | 0.98 | 358.4*** | 0.0128575 |
| BoostD = $C_9 + \beta_9$POP | 150.844*** | 0.99 | 499.52*** | 0.528937 |

Table 3 Part B: Cross-Validation Tests.

| RMSE | *R2* | MAE | Resample | Sample size |
|---|---|---|---|---|
| 219.56 | 1 | 184.56 | - | - |
| | Details of Folds | | | |
| 81.37 | na | 81.37 | Fold 1 | 7 |
| 446.22 | 1 | 339.74 | Fold 2 | 6 |
| 219.56 | 1 | 207.96 | Fold 3 | 6 |
| 202.84 | 1 | 145.95 | Fold 4 | 6 |
| 147.79 | na | 147.79 | Fold 5 | 7 |

Table 3 reveals in the first row of Part A, that excess deaths are significantly related to booster doses, the relationship is significant at the 1 per cent level and the Adjusted R-Squared in the bivariate regression is 0.71. In the second row there is a very similar bivariate regression relationship between excess deaths and total vaccinations, which is significant at the 1 per cent level with a similar Adjusted R-Squared of 0.71. The third row of Table 3, Part A, shows that there is also a similar relationship between excess deaths and the recently vaccinated cohort by State which is also significant at the 1 per cent level and has an Adjusted-R-Squared of 0.69. These first three rows of results suggest that there is a consistent relationship between excess deaths and various vaccination measures across the Australian States.

By contrast, the fourth row in Table 3 reports that there is no significant relationship between excess deaths and the unvaccinated by State. Row 5 in Table 3 reports the results of a bivariate regression of excess deaths by State on reported COVID cases by State. The relationship, like all the previous regressions is positive and significant at the 5 per cent level but the Adjusted R-Squared is reduced to 0.63. Row 6 in Table 3 reports the result of regressing excess deaths by State, on total population by State. This is also significant at the 1 per cent level and has an Adjusted R-squared of 0.66. This result demonstrates the weakness of the analysis, as clearly the more populous States will have more excess deaths, more COVID cases and more vaccinations.

The last 3 rows in Table 3, Part A, report the results of the exploration, via bivariate regressions, of other significant relationships between pairs of variables. Row 7 reports the results of the regression of total vaccinations on total cases by State. The regression is highly significant at the 1 per cent level and the Adjusted R-Squared is 0.98. Similarly, Row 8 in Table 3 explores the relationship between booster doses and total cases. This relationship is also significant at the 1 percent level and has an Adjusted R-Squared of 0.98. Finally, the last row in Table 3, Part A, explores the bivariate relationship between booster doses of the vaccine and population. This is significant at the 1 per cent level and has an Adjusted R-squared of 0.99.

These results are entirely consistent with the correlation matrix and the heat map of correlations presented in

Supplementary Figure 1. Clearly, State population has a large impact on all the variables used in this cross-sectional regression analysis. The concern is that it could be a potential cause of endogeneity, or omitted variable bias, which would reveal itself in a significant relationship between the residuals from the original bivariate regression and the omitted variable. In the final column of Table 3, Part A, I report the F-statistic from the regression of the residuals of the first bivariate regression, reported in the row, with population. The two exceptions to this are reported in rows 6 and 9, in which population is included in the bivariate regression. In row 6, the F-statistic in the final column reports the results of regressing the residuals from the regression on booster doses, whilst in the last row and final column, the F-statistic is the result of regressing the residuals on excess deaths. None of the F statistics in the final column of Part A, are significant, which gives me more confidence in the bivariate regression results.

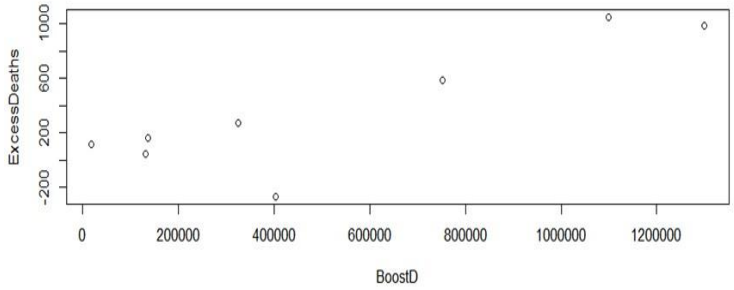
4.2. CROSS-VALIDATION

This section reports the application of this method to the regression of excess deaths on excess vaccine booster doses. I first considered using an R package "cv", Fox and Monette (2025) [15]. Subsequently, as an alternative, I used some GRETL add in code called CvDataSplitter, written by A. Tarassow and available on GITHUB at (https://github.com/atecon/CvDataSplitter, accessed February 8, 2024). In the end, given that there are a limited number of observations, I wrote some R script to produce the output featured in Table 3. Part B.

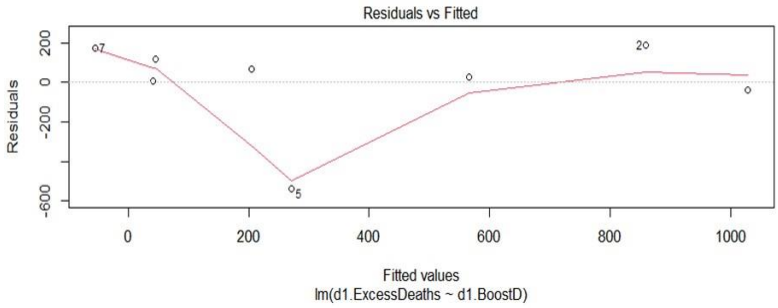
The three metrics, root mean squared error (RMSE), R-squared, and the mean absolute error (MAE), as reported in the second row of Table 3, Part B, under the column headings, show how well the estimated model, performed on the data left out in the folding process. The entries in the bottom 5 rows of Table 3, Part B, under the heading, "details" of the folds, show how the model performed on the data used in each fold and the final column provides details of the sample size. It can be seen that the base model holds up well, the errors are small and the R-squared, a measure of the correlation between the predictions made by the model and the actual observations is 1.

The graphs in Figure 2, show the behaviour of the residuals in the OLS regression of ExcessDeaths on BoostD. These behave well and suggest that the model is well-specified.

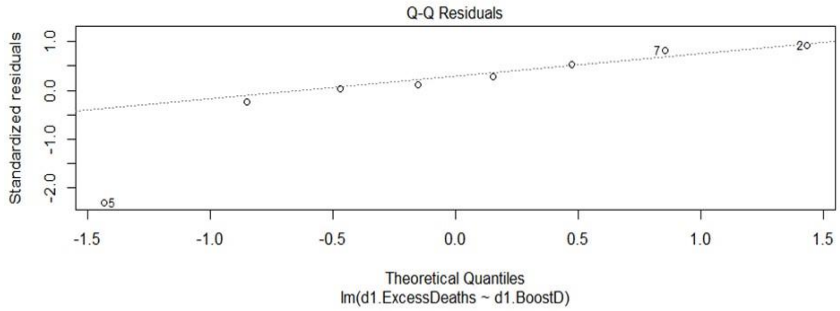
**Figure 2**: Graphs of the data and the behaviour of residuals from the OLS regression of ExcessDeaths on BoostD
Data Plot




Residuals Fit




QQPlot Residuals



## 5. Conclusion

This paper presents the results of an OLS regression of the number of excess deaths, by State and Territory in Australia, as reported in official sources in 2023, on recorded excess booster vaccine doses by State and Territory. The results are quite striking and suggest the existence of a strong regression relationship with significant coefficients and an Adjusted R-squared of 71 percent. Simple cross-validation tests suggest that the regression results are robust, despite the very limited sample size available in this cross-sectional regression. These results match those of a recent study by Allen (2023) on the OECD country experience of excess deaths [1]. They suggest that this topic deserves greater scrutiny given that it is an important public health policy issue with considerable cost implications.

STATEMENTS:
The author gratefully acknowledges the helpful comments of an anonymous reviewer.
The author is not aware of any conflicts of interest.
This study did not benefit from any funding.
The study uses Australian Government data: (https://www.abs.gov.au/articles/measuring-australias-excess-mortality-duringcovid-19-pandemic-until-first-quarter-2023, accessed 9/1/2024), https://www.health.gov.au/our-work/covid-19-vaccines/vaccination-numbers-andstatistics.
Ethics approval was not required for the study.
Author contributions. Not applicable sole author.
Informed consent. Not applicable.

## References


1. Allen, D.E. Excess Deaths and Excess Covid Booster Vaccine Doses – are they related?, *Medical Research Archives*, 2023. Vol 11, Issue 12, https://doi.org/10.18103/mra.v11i12.4841.
2. Gao, F., V. Mallajosyula, P.S. Arunachalam, K. van der Ploeg, M. Manohar, K. Ro¤ ltgen, F. Yang, O. Wirz, R. Hoh, E. Haraguchi, J-Y.' Lee, R. Willis, V. Ramachandiran, J. Li, K. R. Kathuria, C. Li, A. S. Lee, M.M.. Shah, S.B. Sindher, J. Gonzalez, J.D. Altman, T.T. Wang, S.D. Boyd, B. Pulendran, P. Jagannathan, K.C. Nadeau, and M..M. Davis. Spheromers reveal robust T cell responses to the Pfizer/BioNTech vaccine and attenuated peripheral CD8+ T cell responses post SARS-CoV-2 infection, *Immunity* 56, 2023. 864–878, https://doi.org/10.1016/j.immuni.2023.03.005.
3. Krug A, Stevenson J, Høeg T.B. BNT162b2 Vaccine-Associated Myo/Pericarditis in Adolescents: A Stratified Risk-Benefit Analysis, Eur J Clin Invest. 2022;52:e13759. https://doi.org/10.1111/eci.13759
4. Chennupati, K., Bendi S., Gantasala P., Kickloo A., Navigating the Gastrointestinal Implications of Coronavirus 19: Strategies and Lessons Learned, Medical Research Archives, 2024. [online] 12(5). https://doi.org/10.18103/mra.v12i5.5342
5. Pillay J, Gaudet L. Wingert A., Bialy L, Mackie A S, Paterson I.D, Hartling L, Incidence, risk factors, natural history, and hypothesised mechanisms of myocarditis and pericarditis following covid-19 vaccination: living evidence syntheses and review, BMJ 2022; 378: e069445 | doi: 10.1136/bmj-2021-069445
6. Mostert S., Hoogland M., Huibers M., Kaspers G., Excess mortality across countries in the Western World since the COVID-19 pandemic: 'Our World in Data' estimates of January 2020 to December 2022. BMJ Public Health 2024;2:e000282. doi:10.1136/ bmjph-2023-000282
7. Yamamoto, K. Adverse effects of COVID-19 vaccines and measures to prevent them, *Virology Journal*, 2022. 19:100, https://doi.org/10.1186/s12985022-01831-0.
8. Bates, S. Hastie T. and Tibshirani, R. Cross-Validation: What Does It Estimate and How Well Does It Do It? *Journal of the American Statistical Association,* 2023. 1-12. https://doi.org/10.1080/01621459.2023.2197686
9. Zhang, P. Model Selection Via Multifold Cross Validation, *The Annals of Statistics*, 1993. 21, 299–313.
10. Zhang, P. Assessing Prediction Error in Non-parametric Regression, *Scandinavian Journal of Statistics,* 1995. 22, 83–94. [2,4]
11. Hastie, T., Tibshirani, R., and Friedman, J. *The Elements of Statistical Learning* (2nd ed.), New York: Springer. 2009. [2,4]
12. Yousef, W. A. A Leisurely Look at Versions and Variants of the Cross Validation Estimator, 2020. arXiv preprint. arXiv:1907.13413. [2,4]
13. Rosset, S., and Tibshirani, R. J. From Fixed-x to Random-x Regression: Bias-Variance Decompositions, Covariance Penalties, and Prediction Error Estimation, *Journal of the American Statistical Association*, 2020.115, 138–151. [2,6]
14. Wager, S. Cross-Validation, Risk Estimation, and Model Selection: Comment on a Paper by Rosset and Tibshirani, *Journal of the American Statistical Association*, 2020. 115, 157–160. [2,3,5,6]
15. Fox J, Monette G. _cv: Cross-Validation of Regression Models_. R package version 1.1.0, 2024. <https://CRAN.R-project.org/package=cv>.