




## ORIGINAL ARTICLE OPEN ACCESS

# Long Covid Symptom Clusters, Correlates and Predictors in a Highly Vaccinated Australian Population in 2023

Essa Tawfiq<sup>1</sup> | Rosalie Chen<sup>1</sup> | Damian Alexander Honeyman<sup>1</sup>  | Rebecca Dawson<sup>1</sup> | Mohana Kunasekaran<sup>1</sup> | Adriana Notaras<sup>1</sup> | Deepti Gurdasani<sup>1,2,3</sup> | Helen Skouteris<sup>4</sup> | Darshini Ayton<sup>4</sup>  | Chandini Raina MacIntyre<sup>1,5</sup> 

<sup>1</sup>Biosecurity Program, The Kirby Institute, Faculty of Medicine and Health, The University of New South Wales, Sydney, Australia | <sup>2</sup>University of Western Australia Medical School, University of Western Australia, Perth, Australia | <sup>3</sup>William Harvey Research Institute, Queen Mary University of London, London, UK | <sup>4</sup>Health and Social Care Unit, School of Public Health and Preventive Medicine, Monash University in collaboration with Monash Health, Melbourne, Australia | <sup>5</sup>College of Public Service & Community Solutions, and College of Health Solutions, Arizona State University, Tempe, Arizona, USA

**Correspondence:** Deepti Gurdasani ([dgurdasani@protonmail.com](mailto:dgurdasani@protonmail.com)) | Chandini Raina MacIntyre ([r.macintyre@unsw.edu.au](mailto:r.macintyre@unsw.edu.au))

**Received:** 5 July 2024 | **Revised:** 10 April 2025 | **Accepted:** 11 April 2025

**Funding:** E.T., R.C., D.A.H., R.D., M.K., A.N. and D.G. are supported by the Balvi Filantropic Fund. C.R.M. is funded by NHMRC Investigator Grant 2016907. D.A.H. and A.N. are funded through MRFF (grant ID 2017048). The funding sources had no role in the study design, data collection, analysis or interpretation, reporting or publication of this work.

**Keywords:** Australia | Covid-19 | long Covid | Omicron | SARS-CoV-2

## ABSTRACT

**Background:** Limited data exists regarding long Covid burden following Omicron infection in highly vaccinated populations.

**Objective:** To (1) characterise long Covid prevalence and predictors and (2) identify key symptom clusters and their correlates among long Covid patients, during an Omicron-predominant period in a highly vaccinated population.

**Design:** Anonymous, online, cross-sectional survey.

**Setting:** January 2023, Australia.

**Participants:** Residents aged  $\geq 18$  years with self-reported history of test-positive Covid-19.

The main variables studied were socio-demographic characteristics, Covid-19 risk factors, vaccination, infection history and experiences with long Covid.

**Main Outcome Measures:** Long Covid symptoms. Symptom-based clustering was used to identify long Covid symptom clusters and their functional correlates. Predictors of long Covid occurrence and severity were assessed using multivariable logistic regression.

**Results:** Overall, 240/1205 participants (19.9%) reported long Covid. Long Covid risk was significantly higher for women OR 1.71 (95% CI: 1.17–2.51), people with comorbidities 2.19 (95% CI: 1.56–3.08) and those using steroid inhalers for Covid-19 treatment (2.34 [95% CI: 1.29–4.24]). Long-Covid risk increased with increasing Covid-19 infection severity (moderately severe symptoms: 2.23 [95% CI: 1.50–3.30], extremely severe symptoms: 5.80 [95% CI: 3.48–9.66], presented to ED/hospitalised: 7.22 [95% CI: 3.06–17.03]). We found no significant difference in the likelihood of long Covid between the Omicron and pre-Omicron periods, vaccination status and participant age.

We identified two long Covid clusters (pauci-symptomatic,  $n = 170$ , vs. polysymptomatic,  $n = 66$ ). Polysymptomatic cluster membership was associated with worse functioning (impacts on work, moderate activity, emotions and energy). Severity acute infection was strongly predictive of polysymptomatic cluster membership (5.72 [2.04–17.58]). Monoclonal antibody treatment was strongly associated with pauci-symptomatic cluster membership (0.02 [0.00–0.13]).

Essa Tawfiq, Rosalie Chen and Damian Alexander Honeyman are co-first authors.

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2025 The Author(s). *Health Expectations* published by John Wiley & Sons Ltd.

**Discussion:** Our study shows that long Covid is an important health burden in Australia, including during the Omicron era, and identifies several risk factors. We found a subgroup of patients characterised by more symptoms and worse functional outcomes. Our findings can inform policies for protecting vulnerable populations and frameworks for long Covid risk assessment and management.

**Conclusions:** One-in-five people may suffer long Covid after acute Covid-19 infection, with similar risk across age groups. Omicron variants appear not to have a lower risk compared to earlier variants in our study. A cumulative number of symptoms can help triage long Covid patients.

**Patient or Public Contribution:** We did not involve patients or the public in the design of the questionnaire. However, after a soft launch, we revised some survey questions by reviewing early responses from patients and the public.

## 1 | Introduction

Australia experienced one of the largest SARS-CoV-2 epidemics in the Asia-Pacific region between November 2021 and January 2023, largely caused by Omicron [1]. Symptoms that continue after acute SARS-CoV-2 infection are referred to as 'long-COVID', defined by the World Health Organization (WHO) as the continuation or development of new symptoms 3 months after SARS-CoV-2 infection, lasting at least 2 months without other explanation [2]. However, the lack of a global consensus definition and the differences in study designs, including measured symptoms, context (community/hospital), age distribution and vaccination coverage [3], have complicated the estimation of long Covid prevalence. Wide-ranging estimates between 4.7% and 80% have been published between 2021 and 2022 [4], while more recent studies have reported point prevalence estimates of 6%–7% in adults and ~1% in children [5–9]. The global estimated cumulative incidence of long Covid by the end of 2023 was approximately 400 million [10].

Long Covid is a heterogeneous condition with a broad range of symptoms and multisystemic pathology [10, 11]. Over 100 long Covid symptoms have been described [12]. A recent systematic review and meta-analyses found that fatigue, disturbed sleep and breathlessness were highly prevalent among hospitalised, non-hospitalised and mixed cohorts [13]. Clinical diagnosis and management are challenging, due to the variable symptom profile, potentially lengthy symptom duration and lack of diagnostic and referral protocols [10, 14]. Factors, such as hospitalisation, severe acute illness, female gender, older age, lower socioeconomic status and the presence of comorbidities, have been associated with long Covid [15–20]. Whilst symptomatic infection is more likely to cause long Covid, so can asymptomatic infection [21, 22]. Given the association between severe acute illness and long Covid, treatments that reduce the severity of acute Covid-19 infection, such as antivirals, inhaled corticosteroids and monoclonal antibodies [23–25], may modulate long-Covid risk, with one systematic review reporting a protective effect with antivirals, but not monoclonal antibodies [26].

A recent Australian survey found 18.2% of respondents from a highly vaccinated population had ongoing symptoms 90 days after a Covid-19 infection; however, this study was restricted to Western Australian residents [27]. Reduced clinical severity and risk of hospitalisation have been assumed for the Omicron variant compared to the Delta variant, although much of the

data are confounded by vaccination rates, which were higher globally by the end of 2021 when Omicron emerged [28]. Some studies suggest Omicron is less likely to cause long Covid compared to prior variants [29, 30]. The Australian context presents a unique opportunity to characterise long Covid during the Omicron era, as there was relatively low community SARS-CoV-2 transmission before Omicron's emergence, due to border closures and other public health orders [31]. Two-dose vaccination rates in Australians aged 16 years and over were 95% by the end of 2021, when the Omicron wave commenced [32]. We therefore aimed to identify the prevalence of long Covid during an Omicron-dominant period in a highly vaccinated population in Australia and assess the predictors of long-Covid risk. We also aimed to assess if any long Covid symptom-based patient subgroups appeared in the data and to characterise them by phenotypic clusters and functional outcomes.

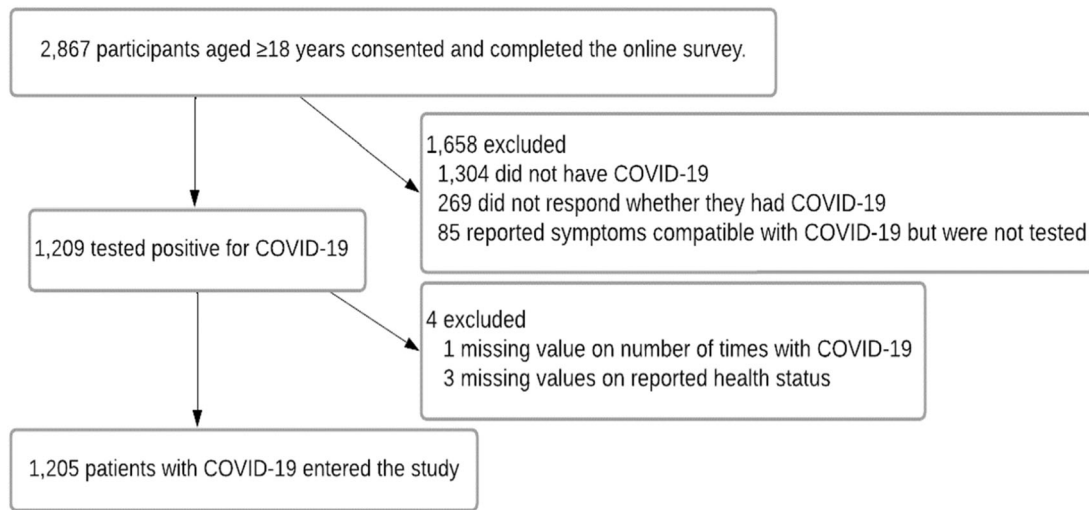
## 2 | Methods

### 2.1 | Study Design and Participants

An anonymous, online, cross-sectional survey of Australian residents aged  $\geq 18$  years was conducted between 9 and 25 January 2023. Eligible subjects were initially identified using a global market research company (Dynata) [33]. Data on socio-demographic characteristics, Covid-19 risk factors, pre-existing health status, vaccination and infection history were collected using the web-based survey platform Qualtrics (for further details, see [Supporting Information](#), Methods—Additional Information). Participants were included in this study if they reported testing positive for SARS-CoV-2 using a polymerase chain reaction (PCR) test and/or rapid antigen test (RAT) (Figure 1).

### 2.2 | Outcomes

The outcome of 'long COVID' was based on the WHO Delphi definition [2]. Respondents were asked whether they had 'symptoms that occurred *within* 3 months of the onset of Covid-19 and lasted for at least 2 months'. This differs from the WHO definition in that, theoretically, this may include participants who had continuous symptoms for 2 months post-Covid that resolved within 3 months (i.e., participants who did not have symptoms at 3 months). For those who answered 'yes', data on



**FIGURE 1** | Study population and inclusion and exclusion criteria.

the type and duration of a list of pre-specified long-Covid symptoms was collected, with a textbox field for other symptoms. We also asked about changes in self-reported health status (Excellent, Very good, Good, Fair and Poor) pre- and post-acute Covid-19 infection based on participant recall at the time of survey. The percentage change in health rating was calculated as the difference in number of patients in each rating category pre- and post-Covid-19, divided by the number of patients in the pre-Covid-19 rating category.

## 2.3 | Predictors

We examined a pre-specified list of predictors based on evidence from the literature from studies examining predictors for long Covid and/or acute Covid-19 severity (as many studies have shown that acute Covid-19 severity correlates with risk of long Covid) [14, 34] (Figure S1). We considered age, gender, socio-economic status, pre-existing health status, comorbidities, dominant variant at the time of first infection, severity of acute infection, number of infections and vaccination as predictors, given these had been associated with long Covid risk in previous studies [35–39]. We also included treatments (oral antivirals, monoclonal antibodies, steroid inhalers and other medications) for acute-Covid-19, as acute Covid-19 severity is known to be associated with long Covid [40]. We also examined the association of mask use with long Covid, given viral dosage has been hypothesised to be associated with acute Covid-19 severity (Figure S1) [41]. We used employment status, house ownership, whether the participant was born in Australia, education level, type of accommodation and family size as possible socio-economic indicators (see Supporting Information for details). We note that all exposures collected through the questionnaire were based on participant recall at the point of survey, and characteristics such as employment would reflect participant status at the point of survey rather than infection. For information on re-infection, we only collected the total number of infections at the point of survey, which may not reflect the number of infections before the development of long Covid, leading to misclassification, and conservative estimates of results. We classified the variant causing infection based on

the first infection, which may or may not have been the cause of long Covid, which may again lead to misclassification, and conservative estimates.

### 2.3.1 | Statistical Analysis

We first conducted a descriptive analysis of the baseline characteristics of participants who reported infection, with and without long Covid. We used the  $\chi^2$  test to evaluate the statistical significance of differences in baseline characteristics (excluding cells with counts of < 5) (Supporting Information, Methods—Additional Information). To examine the change in health following the development of long Covid, we first calculated the proportion of participants who had worsening health, or improving health, or whose health was stable, post-long Covid. We then carried out a one-sample test for difference in proportion from 0 for the proportion with worsening health. This assesses whether the proportion of those with worsening health was statistically significantly different from zero.

To examine predictors of long COVID, we used multivariate logistic regression. We first examined the pre-specified list of predictors based on the literature (Figure S1), as outlined above (Model 1). To avoid multicollinearity and overfitting, we also carried out backwards selection, iteratively removing the variable with the highest  $p$  value, until all variables included had a  $p$  value of < 0.10, and presented these in a second model (Model 2).

Variables included in the full model were age (continuous), gender (male/female), employment status at the time of survey (employed/unemployed), home ownership (yes/no), being born in Australia (yes/no), education level (primary school, high school, TAFE or higher education), type of accommodation (house vs. unit/townhouse/other), self-reported health status before Covid-19 infection (poor vs. good), mask use (yes/no), vaccination status (0–2 doses vs.  $\geq 3$  doses), presence of comorbidities (at least one comorbidity vs. no comorbidity), infection with pre-Omicron versus Omicron variant (first positive Covid-19 test before December 2021 vs. from December

2021 onwards) [30], number of Covid-19 infections (ordinal), severity of acute Covid-19 infection (mild, moderately severe symptoms, extremely severe symptoms and presented to emergency department) (Table S1), and Covid-19 treatment with antivirals (yes/no), steroid inhalers (yes/no), monoclonal antibodies (yes/no) and other drugs taken for Covid-19 treatment (no treatment with these drugs/treated with hydroxy-chloroquine, ivermectin, fluvoxamine or fluoxetine). When assessing steroid inhalers as a treatment for Covid-19, patients with pre-existing respiratory illness (asthma, chronic bronchitis or COPD) were not considered as having received this specifically as treatment for acute COVID, as we were interested in evaluating these as a treatment specifically for acute Covid-19 infection. We note that considering this group as not having received these specifically for long Covid may lead to more conservative results, biasing our effect estimates towards the null, as some in this group may have received these specifically for acute infection. Categories with small cell counts were excluded from regression analysis (see [Supporting Information, Methods—Additional Information](#)).

We examined the assumption of linearity on the logit scale of continuous or ordinal variables using the Box–Tidwell test [42] and found that this was violated for the number of pre-existing comorbidities when coded as an ordinal variable. This was not violated for any of the other continuous or ordinal variables. We recoded pre-existing comorbidities as a binary category, as detailed above, with no comorbidity as the reference. Given the multiple predictors tested for, we used a Bonferroni-corrected threshold for  $\alpha$ . To maintain an  $\alpha$  of 0.05, we used a two-way  $p$  value threshold of  $< 0.003$  for statistical significance (corrected for 18 tests). We note that, given the conservative nature of Bonferroni correction, our results are likely to be conservative, strictly controlling the false discovery rate at the cost of reduced statistical power. To provide a less conservative comparator, we also present the Benjamini–Hochberg correction [43] using FDR of 5% to calculate adjusted Benjamini–Hochberg  $p$  values based on ranked  $p$  values for 18 predictors of long Covid. Those below 0.05 were considered statistically significant in this exploratory analysis.

We also conducted sensitivity analyses to examine the effect of re-categorising variables differently: vaccination status (unvaccinated, 1–2 doses and  $\geq 3$  doses), number of COVID-19 infections (once, twice and three or more times) and age as a categorical variable (18–50, 51–64 and  $\geq 65$  years).

Given that the severity of acute Covid-19 and the presence of comorbidities are strongly associated both with long Covid severity (Figure S1), and treatment eligibility, we assessed the potential for residual confounding of the association between treatments and long Covid by these variables. To better assess potential residual confounding of the association between treatments and long Covid by acute Covid severity, and pre-existing co-morbidities, we first examined the distribution of Covid-19 treatments across strata of acute Covid-19 severity and comorbidities. We then conducted subgroup analyses, examining the relationship between Covid-19 treatments and long Covid within strata of acute Covid-19 severity (mild, moderately severe, and extremely severe or presented to ED/hospitalised) and comorbidities (no comorbidity, 1 comorbidity and  $\geq 2$  comorbidities).

We also examined whether symptoms of long Covid varied between the participants who reported first infections during the pre-omicron and omicron period. We used the  $\chi^2$  test to examine differences in cell counts for those with and without each symptom.

Our sample of 1205 participants with reported SARS-CoV-2 infection allows estimation of long Covid prevalence at a true prevalence of 10% with 2% precision. Our sample of 240 and 965 participants, with and without long Covid, respectively, provides 80% power to detect a proportional difference of 3.4% for exposures present at 20% prevalence in the non-long-Covid group with an  $\alpha$  error of 0.05.

## 2.4 | Clustering Analysis

To identify subgroups of participants by symptoms, we carried out clustering analysis across all symptoms, excluding those with low counts ( $< 5$  participants). As different methods of clustering can result in different subgroups, and there is no single method that is the gold standard, we first conducted an exploratory analysis of two clustering algorithms, Partitioning Around Medoids (PAM) and Agglomerative Hierarchical clustering, both widely used approaches for unsupervised clustering of clinical data [44]. Clustering was performed in ‘R’ (version 2023.03.1+446), using the ‘cluster’ and ‘stats’ packages [45, 46]. Participants were clustered on 19 unique long Covid symptoms recorded, each coded as a binary variable. A Manhattan distance matrix was used to quantify dissimilarity between participants based on symptoms and assign them to clusters. We note that for binary data, the Manhattan distance is equivalent to Hamming distance, an effective measure for comparing two equal-length binary strings [47].

Average Silhouette coefficient score, a measure of intra-cluster cohesion and inter-cluster separation [48] with a range from  $-1$  to  $+1$ , was calculated to identify the optimal number of clusters for both methods. A greater positive score indicates that a datapoint is similar to datapoints in the same cluster and different to datapoints in another cluster; a greater negative score indicates likely misclassification. Cluster stability for each method was assessed with Jaccard’s coefficient using bootstrapping with 100 resamples to evaluate the extent to which random resampling altered cluster membership. We also assessed the similarity of clusters between the two methods with Jaccard’s coefficient. This compares the similarity of clusters identified by calculating a ratio of their intersection to the union size. A value of 0 indicates completely different sets, and 1 indicates completely identical sets.

Participant characteristics, including demographic and socioeconomic variables, vaccination status, severity of Covid-19 infection and treatment types, were summarised for each cluster, with statistical differences evaluated using Pearson’s  $\chi^2$  test for categorical data (excluding cell counts  $< 5$ ) and Kruskal–Wallis test for ordinal and continuous data.

To examine the functional and clinical significance of symptom clusters, we also examined functional correlates of cluster membership, including physical limitation, emotional status,



social impacts, occupational consequences and self-ratings of health post-COVID. Pearson's  $\chi^2$  test for categorical data (excluding cell counts < 5) and Kruskal–Wallis test for ordinal and continuous data were used to compare baseline characteristics and functional correlates between clusters. A multiple logistic regression model was fit using the same predictors as used in the multivariate model for long Covid risk, using cluster membership as the dependent variable, to identify predictors of cluster membership. Consistent with our statistical analysis of predictors of long Covid, we present two models for the prediction of cluster membership, one with all pre-specified potential predictors included, and one using backwards regression, using a threshold of  $p < 0.10$ . Results are presented in tables, butterfly charts, forest and alluvial plots, created using R ('plotrix', 'metafor' and 'ggalluvial' packages) and MATLAB [49].

## 2.5 | Ethical Considerations

This study was approved by the UNSW Human Research Ethics Committee (approval number HC220737).

## 3 | Results

One in five (240/1205 [19.9%]) participants in our study reported long Covid symptoms. Statistically significant differences in baseline characteristics (Table 1) were observed between those with and without self-reported long Covid, with long-Covid patients being younger adults of working age ( $\chi^2 = 13.98$ ,  $p = 0.001$ ), disproportionately female ( $\chi^2 = 17.60$ ,  $p = 0.001$ ) and more likely to report reinfection ( $\chi^2 = 20.65$ ,  $p < 0.001$ ) and pre-existing comorbidities ( $\chi^2 = 18.17$ ,  $p < 0.001$ ). However, there was no statistically significant difference by vaccination status and whether first infection occurred during or before the Omicron-dominant era (Table 1). Consistent with pre-existing conditions, and severe acute COVID-19 being overrepresented among those with self-reported long COVID, this group were also more likely to report molnupiravir ( $\chi^2 = 11.39$ ,  $p = 0.001$ ), steroid inhaler ( $\chi^2 = 32.13$ ,  $p < 0.001$ ) and monoclonal antibody ( $\chi^2 = 50.56$ ,  $p < 0.001$ ) use compared to those without long Covid.

Of the 240 participants with long Covid, a decline in health status was seen post-Covid (Figure S3 and Table S2), with a statistically significant proportion (52%) reporting a decline in health ( $p$  value < 0.0001), only 1% reporting an improvement and 47% stable health post-Covid (Figure S3). Tiredness was the most reported long Covid symptom (18.5%) followed by shortness of breath (12.0%) (Figure S4).

On multivariate regression (Figure 2 and Tables S3 and S4), severe acute Covid-19 infection was a strong predictor of long Covid, with double the risk (2.23 [95% CI: 1.50–3.30],  $p = 0.0001$ ) for moderately severe symptoms, nearly six times for extremely severe symptoms (5.80 [95% CI: 3.48–9.66],  $p = 0.0001$ ), and seven times for presenting to ED (7.22 [95% CI: 3.06–17.03],  $p = 0.0001$ ), compared to mild illness. Having at least one pre-existing comorbidity was associated with twice the odds of long Covid (2.19 [95% CI: 1.56–3.08],  $p = 0.0001$ ). Steroid inhaler and monoclonal antibody use for the treatment of

Covid-19 were associated with higher risk (2.34 [95% CI: 1.29–4.24],  $p = 0.005$  and 3.24 [95% CI: 1.74–6.02],  $p = 0.001$ , respectively). Age, female gender and steroid inhaler use were nominally associated with higher risk of long Covid. However, these differences were not statistically significant after Bonferroni multiple correction. However, female gender and steroid inhaler use were statistically significant when using the less conservative Benjamini–Hochberg correction (Table S3). Model 1 (the full model with all pre-specified variables) and Model 2 (limited model following backwards stepwise regression) both showed consistent effect estimates (Tables S3 and S4). Sensitivity analyses with re-categorised variables (vaccination status, number of times with Covid-19 infection and age) produced similar results as Model 1 (Tables S5, S6 and S7).

Exploratory analyses examining the role of confounding between the association between monoclonal antibody use and long Covid risk showed that monoclonal antibody use was disproportionately overrepresented in those with severe symptoms, or emergency department attendance or hospitalisation, compared to those with mild or moderate symptoms (Table S8). Similarly, monoclonal antibody use was also overrepresented among those with one or more comorbidities (Table S9). Sensitivity analyses within strata of severity showed a statistically significant strong positive association between monoclonal antibody use and long Covid within the mild strata of acute Covid-19 severity ( $p = 0.001$ ) (Table S10). However, within strata of comorbidities, no statistically significant association was observed between monoclonal antibody use and long Covid (Table S11), although the association was nominally significant (not attaining the threshold for Bonferroni correction) among participants with two or more comorbidities.

We found no material difference between long-Covid symptoms for participants reporting their first infection in the pre-Omicron versus Omicron era, except for palpitations, which were more common in those reporting infection in the pre-Omicron period (6.3%) compared to those in the Omicron period (1.7%), where the difference was statistically significant ( $p < 0.003$ ) (Table S12).

## 3.1 | Clustering Analysis

Both clustering methods identified two clusters as the optimal number of clusters (Figures S1 and S5), with a Jaccard's coefficient of 0.86 between methods, suggesting high concordance between the two methods (Figure S6). The Jaccard's coefficient for the clusters within each method was 1.00. As the representation of clusters and functional correlates of these clusters was similar across the two methods, we primarily present results from PAM clustering here (Tables S13–S14 and Figure S7). Results from hierarchical clustering were consistent with this and can be found in Supporting Information (Figures S8–S11 and Tables S15–S17).

For PAM clustering, 236 participants with long Covid were included, with 170 participants assigned to one cluster and 66 to another cluster, with an average silhouette width of 0.39 (Figure S5). We designate these clusters broadly as paucisymptomatic and polysymptomatic, as the primary difference

**TABLE 1** | Characteristics of participants who had Covid-19 at time of survey.

	<b>Patients with long Covid <i>n</i> = 240 (19.9%)</b>	<b>Patients without long Covid <i>n</i> = 965 (80.1%)</b>	<b>Total number of patients <i>n</i> = 1205 (100%)</b>	<b><i>p</i> value</b>
Age				0.001
18–50 years	179 (22.7%)	611 (77.3%)	790 (100%)	
51–64 years	40 (17.9%)	183 (82.1%)	223 (100%)	
≥ 65 years	21 (10.9%)	171 (89.1%)	192 (100%)	
Gender				0.001
Male	65 (15.0%)	367 (85.0%)	432 (100%)	
Female	170 (22.3%)	594 (77.7%)	764 (100%)	
Non-binary	< 5	< 5	6 (100%)	
Did not state	< 5	< 5	< 5	
Employment				0.006
Employed	178 (22.1%)	626 (77.9%)	804 (100%)	
Unemployed	62 (15.5%)	339 (84.5%)	401 (100%)	
State/Territory				0.20
Australian Capital Territory	38 (20.7%)	146 (79.3%)	184 (100%)	
New South Wales	30 (20.5%)	116 (79.5%)	146 (100%)	
Northern Territory	15 (22.7%)	51 (77.3%)	66 (100%)	
Queensland	29 (18.2%)	130 (81.8%)	159 (100%)	
South Australia	28 (17.3%)	134 (82.7%)	162 (100%)	
Tasmania	28 (23.7%)	90 (76.3%)	118 (100%)	
Victoria	35 (20.2%)	138 (79.8%)	173 (100%)	
Western Australia	37 (18.8%)	160 (81.2%)	197 (100%)	
Vaccination status				0.19
Unvaccinated	9 (14.1%)	55 (85.9%)	64 (100%)	
1–2 dose(s)	58 (24.2%)	182 (75.8%)	240 (100%)	
3+ doses	172 (19.3%)	720 (80.7%)	892 (100%)	
Unsure whether received vaccines	1 (11.1%)	8 (88.9%)	9 (100%)	
Pre-existing comorbidity*				< 0.001
Yes	168 (24.1%)	529 (75.9%)	697 (100%)	
No	72 (14.2%)	436 (85.8%)	508 (100%)	
Omicron as the cause of illness				0.38
Yes	188 (19.4%)	780 (80.6%)	968 (100%)	
No	52 (21.9%)	185 (78.1%)	237 (100%)	
Number of times with Covid-19				< 0.001
1 time	180 (17.9%)	828 (82.1%)	1008 (100%)	
2 times	45 (27.8%)	117 (72.2%)	162 (100%)	
≥ 3 times	15 (42.9%)	20 (57.1%)	35 (100%)	
Severity of acute Covid-19				< 0.001
Mild	39 (9.2%)	385 (90.8%)	424 (100%)	
Moderately severe	131 (20.9%)	497 (79.1%)	628 (100%)	
Extremely severe	52 (42.6%)	70 (57.4%)	122 (100%)	
Presented to ED/hospitalised	18 (58.1%)	13 (41.9%)	31 (100%)	
Antiviral medications				

(Continues)

TABLE 1 | (Continued)

		Patients with long Covid <i>n</i> = 240 (19.9%)	Patients without long Covid <i>n</i> = 965 (80.1%)	Total number of patients <i>n</i> = 1205 (100%)	<i>p</i> value
Molnupiravir used	Yes	19 (38.8%)	30 (61.2%)	49 (100%)	0.001
	No	221 (19.1%)	935 (80.9%)	1156 (100%)	
Paxlovid used	Yes	22 (28.2%)	56 (71.8%)	78 (100%)	0.06
	No	218 (19.3%)	909 (80.7%)	1127 (100%)	
Remdesivir used	Yes	7 (29.2%)	17 (70.8%)	24 (100%)	0.25
	No	233 (19.7%)	948 (80.3%)	1181 (100%)	
Steroid inhalers for Covid-19 treatment					< 0.001
Yes		32 (46.4%)	37 (53.6%)	69 (100%)	
No		208 (18.3%)	928 (81.7%)	1136 (100%)	
Monoclonal antibodies for Covid-19 treatment					< 0.001
Yes		37 (52.9%)	33 (47.1%)	70 (100%)	
No		203 (17.9%)	932 (82.1%)	1135 (100%)	

Note: The Bonferroni corrected *p* value was 0.003, based on a 0.05  $\alpha$  level and 18 predictors.

\*See Figure S2 for a detailed breakdown of comorbidities.

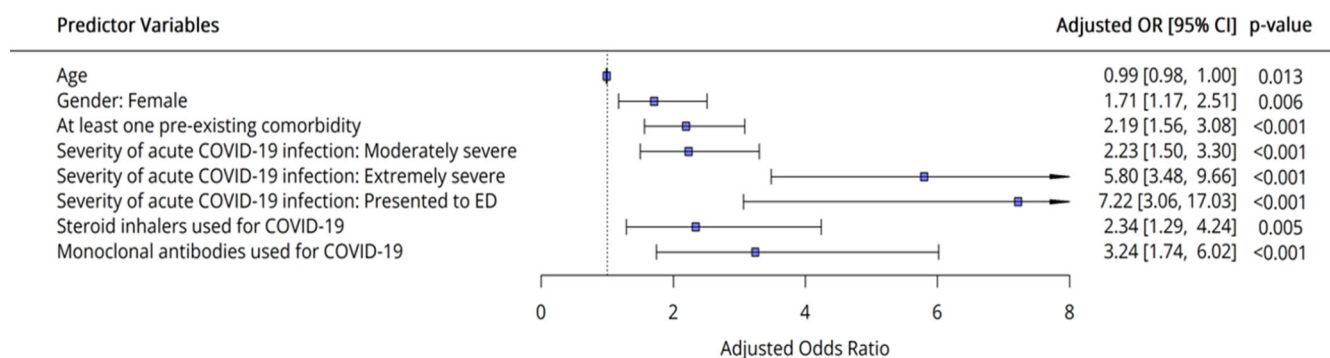


FIGURE 2 | Predictors of long Covid—multivariable regression model (Model 2).

between them appears to be number of symptoms reported by participants (median [IQR] symptom count: 3 [1–4] vs. 9 [7–11],  $H = 137.13$ ,  $p = 0.001$ ) (Figure S7). Both clusters had a median (IQR) symptom duration of 3 [2–5] months.

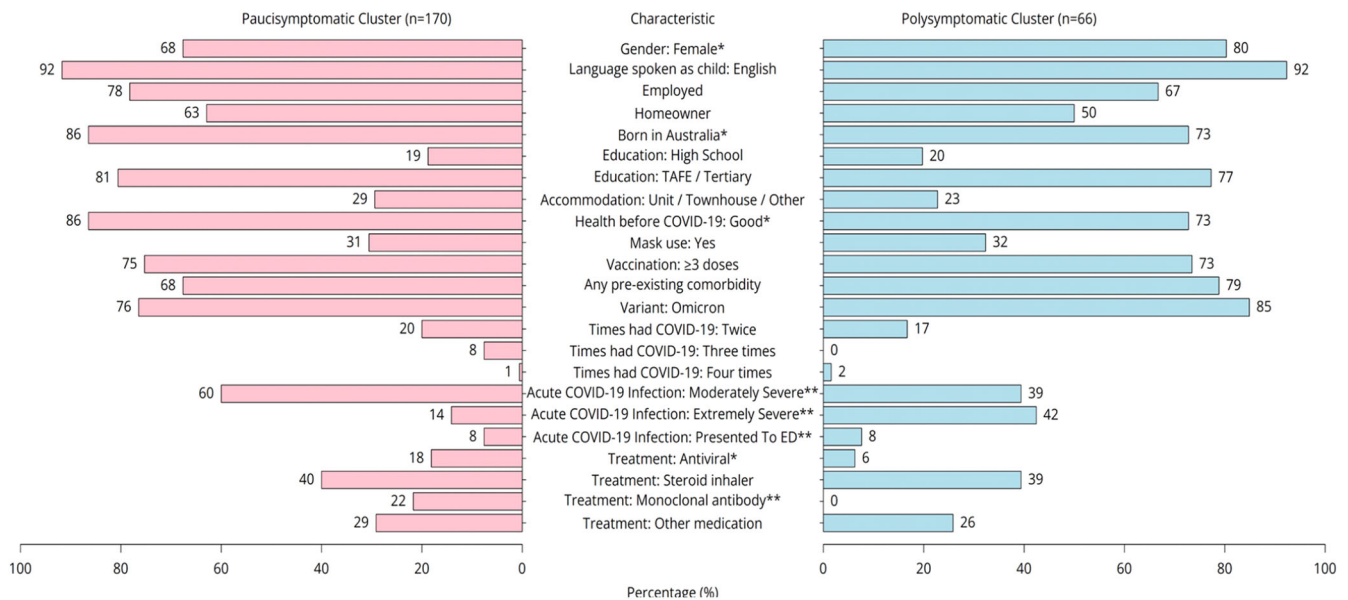
On comparing baseline characteristics, the polysymptomatic cluster had a higher proportion of female participants ( $\chi^2 = 4.91$ ,  $p = 0.03$ ), although this was only nominally significant, and did not attain statistical significance after multiple testing correction. The polysymptomatic cluster also had more participants with severe acute infection ( $\chi^2 = 22.73$ ,  $p = 0.001$ ) and were less likely to have received monoclonal antibody treatment ( $\chi^2 = 15.43$ ,  $p = 0.001$ ) (Figure 3).

In multivariate analysis of predictors of polysymptomatic cluster membership, having Severity of acute infection (extremely severe symptoms in acute infection) was significantly associated with cluster membership (OR = 5.72, 95% CI: 2.04–17.58,  $p = 0.001$ ), and monoclonal antibody treatment was significantly associated with pauci-symptomatic cluster membership

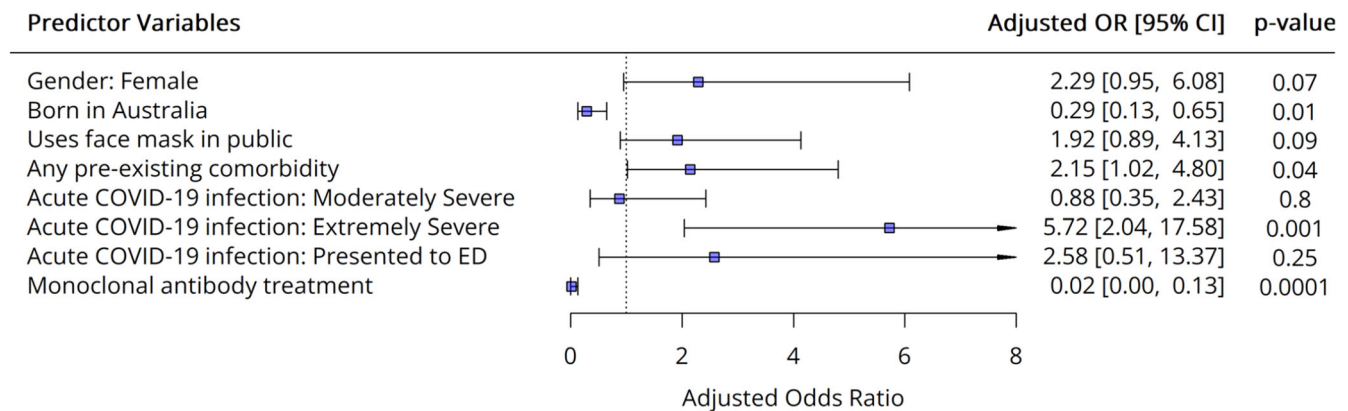
(OR = 0.02, 95% CI: 0.00–0.13,  $p = 0.0001$ ) (Figure 4). Female gender became non-significant in multivariate analysis.

Regarding functional outcomes, polysymptomatic cluster membership was associated with significantly poorer physical outcomes (limitations with moderate activities [ $\chi^2 = 14.12$ ,  $p = 0.001$ ] and climbing stairs [ $\chi^2 = 9.59$ ,  $p = 0.002$ ]), work-related impacts (more likely to work less carefully [ $\chi^2 = 9.03$ ,  $p = 0.003$ ], accomplish less work [ $\chi^2 = 13.0$ ,  $p = 0.001$ ] and report pain interfering with work [ $\chi^2 = 7.53$ ,  $p = 0.006$ ]) (Figure 5A), emotional status (less likely to feel calm, peaceful [ $\chi^2 = 12.66$ ,  $p = 0.001$ ] or energetic [ $\chi^2 = 21$ ,  $p = 0.001$ ] during the last 4 weeks) and social impacts (more likely to report physical or emotional problems interfering with social activities [ $\chi^2 = 4.99$ ,  $p = 0.03$ ]) (Figure 5B), although this was nominally significant and did not attain statistical significance after multiple testing correction.

Polysymptomatic cluster membership was associated with poor reported health at baseline, with a higher proportion of



**FIGURE 3** | Characteristics of long Covid participants belonging to the pauci-symptomatic and polysymptomatic clusters at the time of survey. Kruskal–Wallis test was applied to non-normally distributed continuous variables and ordinal variables. Pearson’s  $\chi^2$  was applied to categorical predictor variables. Categories with cell counts less than 5 were excluded from analysis, which included ‘Gender: Non-binary’. For a detailed frequency table, see Table S13. \* $p$  value < 0.05, \*\* $p$  value < 0.001.



**FIGURE 4** | Forest plot showing adjusted odds ratios and 95% confidence interval for predictors of polysymptomatic cluster membership (PAM method). Blue squares indicate the odds ratio, and horizontal lines depict the 95% confidence interval.

participants reporting their pre-infection health status as ‘Poor’ (18/66 [27.27%]), compared to those in the pauci-symptomatic cluster (23/170 [13.53%]) (Table S13). Those in the polysymptomatic cluster were also more likely to report worsening in health post-infection compared to baseline ( $\chi^2 = 5.50$ ,  $p = 0.02$ ) (Figure 5C) and self-report poorer health status post-infection ( $\chi^2 = 25.28$ ,  $p = 0.001$ ) (Figure 5D).

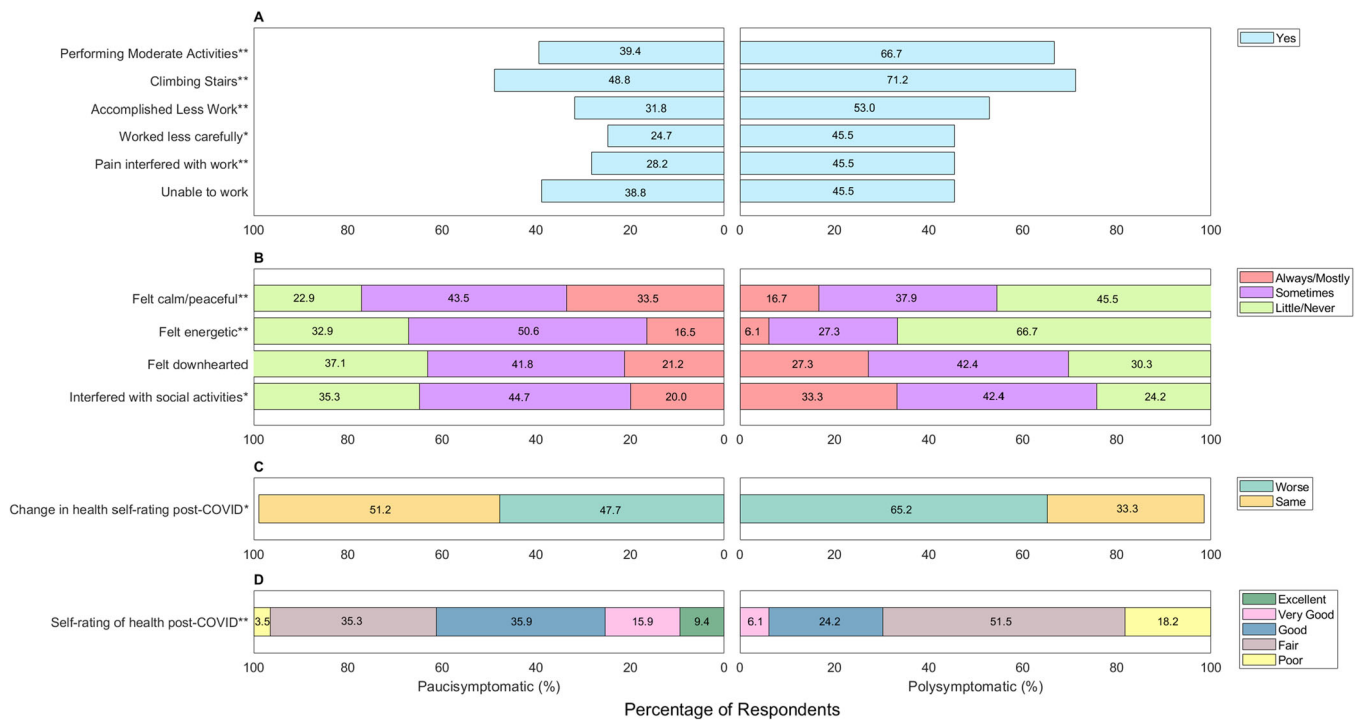
#### 4 | Discussion

Long Covid was found to impact adults of all age groups in our survey, with a prevalence of nearly 20% in a highly vaccinated population. These findings are consistent with other studies [17, 31], adding to a growing evidence base that suggests the prevalence of long Covid post-Omicron infection is higher than previously thought and affects all age groups. We found that the Omicron variant was not significantly different to prior variants

in causing long Covid, despite previous studies suggesting that the Omicron variant is associated with a decreased long-Covid risk compared to the Delta variant, even at 6 months post-vaccination [30]. Although Omicron is believed to be associated with less severe acute infection [29], our findings support evidence that its health impacts may still be substantial, particularly in the form of long Covid, and should not be neglected, even as countries downgrade their severity ratings of SARS-CoV-2 infection.

Consistent with previous studies, we also found that female gender, re-infection, pre-existing conditions and more severe acute infection were associated with developing long Covid [16, 50, 51]. Women have a fast-acting, innate and adaptive immune system, which usually protects from initial disease but also increases the risk of chronic immune-related disorders [52], and they may also have higher exposure to the virus due to caring roles. While abundant research supports that vaccination is





**FIGURE 5 |** (A–D) Functional correlates and post-COVID outcomes of long COVID participants belonging to the pauci-symptomatic and polysymptomatic clusters. Kruskal–Wallis test was applied to non-normally distributed continuous variables and ordinal variables. Pearson's  $\chi^2$  was applied to categorical predictor variables. Categories with cell counts less than 5 were excluded from analysis, which included 'Change in health self-rating post-COVID: Improved'. For a detailed frequency table, see Table S14. \* $p$  value < 0.05, \*\* $p$  value < 0.001.

associated with reduced long Covid severity and social, personal and work impacts, we did not find an association with vaccination. This may be due to very few participants being unvaccinated, the likely lengthy delays since last booster/vaccination, or residual confounding (e.g., those with pre-existing conditions are more likely to be vaccinated and to also get long Covid). Most people who reported three vaccination doses would likely have received their last dose at least 6 months before the survey, and a further booster was not yet available at the time of the survey. Immunity to boosters wanes rapidly, even after a third or fourth dose [53]. More research is therefore required to ascertain the relationship between the recency of vaccination and matching of boosters to circulating variants and long-Covid risk. This is relevant to Australia, where in September 2023, further boosters were only recommended for people aged 75 years and older. For people aged 65–74 years, a booster can be considered 6 months after the last dose or infection, and for people < 65 years, if they are severely immunosuppressed [54]. Our research suggests adults of any age are at risk of long Covid, which should be considered in vaccination policy for all ages.

Antiviral medication has also been found to reduce long Covid risk in unvaccinated and vaccinated cases [55, 56]. However, we found that antivirals were not associated with lower long-Covid risk, possibly because few participants received antivirals, thus rendering the study underpowered for this predictor, or possibly due to residual confounding, as we discuss subsequently. Monoclonal antibodies were found to be associated with an increased risk of long Covid. This is likely due to residual confounding by the presence of health conditions and acute Covid-19 severity, which was associated both with antiviral use,

monoclonal antibody use and a higher probability of long Covid. Stratifying by the number of comorbidities rendered the association between monoclonal antibodies and long Covid non-significant (with effect size reduced), although this may also be partly due to reduced statistical power due to smaller numbers within strata.

The exploratory clustering analysis in this study is novel as it assesses phenotypic differences in long Covid patients and their associated baseline characteristics and functional outcomes in a highly vaccinated Australian cohort predominantly infected with Omicron. We identified robust, reproducible and clinically significant clusters of long Covid symptoms. The strong association of the polysymptomatic cluster with more long Covid symptoms and poorer functional outcomes suggests a more severe and multisystemic symptom profile in a smaller subgroup of long Covid patients, consistent with other studies [57]. We show that monoclonal antibody use predicts pauci-symptomatic cluster membership, suggesting a potentially moderating impact on long Covid symptoms, while more severe acute Covid-19 disease predicts polysymptomatic cluster membership, which is also associated with poorer physical, emotional, social and occupational impacts. Whichever subgroup one belongs to, therefore, may be determined by a combination of factors, including the severity of acute infection and access to treatment. Understanding these subgroups and their predictors may help clinicians to prioritise patients for diagnostics, specialist referral and more aggressive management.

Our study has several limitations. Our survey may not be generally representative of the Australian population, particularly in terms of geographical location and gender. Most respondents

( $n = 197$  [16.3%]) were from Western Australia (the fourth-most populous state/territory), and New South Wales had 146 respondents (12.1%), despite accounting for 31.4% of Australia's population [58]. Participants were mostly female ( $n = 764$  [63.4%]). However, the age distribution and median age of participants (41.0 vs. 38.5 years) and proportion of participants working in healthcare (15.9%) were similar to the general population [58, 59]. Participants were also asked to describe their gender; thus, we could not assess the relationship between sex and long Covid. However, many studies have also reported on the relationship between gender and long Covid [17, 28]. This likely reflects known differences in immunological responses between males and females, with stronger innate immune responses and greater autoimmunity reported in females, influencing the variability in the incidence of autoimmune diseases between sexes [52]. Next, our long Covid definition differs from the Delphi consensus definition, due to the slightly different phrasing of the question, which may lead to inclusion of participants who had continuous symptoms for 2 months post-Covid that resolved within 3 months (i.e., participants who did not have symptoms at 3 months). While symptoms lasting for 2 months are likely functionally significant to the individual, this may have inflated our estimate of long Covid prevalence relative to the Delphi definition, which requires symptoms to be present at the 3-month point. Additionally, questions relating to long Covid were asked only of patients who reported testing positive for SARS-CoV-2, thus, baseline rates of symptoms could not be determined for those who did not test positive. Non-response rates also could not be calculated, as Dynata does not provide the total number of survey recipients. Furthermore, recall bias may be present, as patients' SARS-CoV-2 infections and symptoms may have happened many months ago. Although we used logistic regression and sensitivity analyses to adjust and assess for confounding, our results could have been affected by residual confounding due to unmeasured confounders, such as time since last Covid-19 vaccination and acute infection and measurement error in measured confounders. Some data (e.g., median duration of symptoms) were only collected for those who reported having persistent symptoms, so we could not conduct comparisons for such variables between those with and without long Covid. As patient characteristics (e.g., employment) were collected at the point of survey, this may lead to reverse association bias, as characteristics may have been measured following the development of long Covid and may be caused by long Covid, rather than the reverse. We did not collect the dates of each repeat infection, thus, the total number of infections may not indicate the cumulative number of infections before developing long Covid, as long Covid could have developed before the most recent infection. It is also possible that a patient's first and any subsequent infection(s) were caused by different SARS-CoV-2 variants. These limitations are expected to have led to misclassification bias, which would have biased results towards the null. There is also a possibility of recall bias for exposures, given that exposures and health status were recalled following infection and/or development of long Covid, which may potentially bias results away from the null. Regarding the clustering analysis based on long Covid symptoms experienced, it is possible that the milder symptom profile in the paucisymptomatic cluster may be a result of treatment received, as we found a strong association with monoclonal antibody treatment. We note that while the pauci-symptomatic cluster is

strongly associated with treatment, causality cannot be inferred, given the cross-sectional and observational nature of our survey. Further research is required to elucidate the effect of multiple infections, possibly due to multiple variants, on long Covid risk and severity, such as investigating whether multiple infections impede recovery from long Covid caused by an earlier infection or if subsequent infections indeed increase long Covid risk. Lastly, as the study is observational, causality between the predictors and the outcome of long Covid cannot be examined.

Overall, our study shows that long Covid is an important health burden in Australia and identifies several potential risk factors for its occurrence and severity. Our findings can inform diagnosis, referral and management of patients based on the cumulative number of symptoms, particularly if combined with better organ system-based pathological and diagnostic protocols [60]. These findings could also inform policies to protect vulnerable populations, including minimising community transmission, ensuring adequate vaccination coverage and treatment access, particularly for severe acute Covid-19 infection, producing decision-support tools for assessing individual long Covid risk and guidelines for long Covid management. Further research is warranted to examine the impact of more recent vaccination, particularly in adults of working age, and multiple SARS-CoV-2 infections on long Covid outcomes.

## 5 | Conclusion

This study provides further evidence that long Covid presents an important health burden for Australian Covid-19 survivors of all ages. The Omicron variant appears as likely to cause long Covid as prior variants. We identified several risk factors for long Covid occurrence and severity and found a smaller subgroup of patients who appear to experience a more severe, polysymptomatic profile. Public health policy should support diagnostic, preventive and therapeutic measures for reducing long Covid risk and severity.

---

### Author Contributions

**Essa Tawfiq:** methodology, writing – original draft, writing – review and editing, formal analysis, investigation, data curation, visualization. **Rosalie Chen:** methodology, investigation, visualization, formal analysis, writing – review and editing, writing – original draft, data curation. **Damian Alexander Honeyman:** writing – original draft, writing – review and editing, methodology, formal analysis, investigation, visualization, data curation. **Rebecca Dawson:** methodology, formal analysis, investigation. **Mohana Kunasekaran:** writing – review and editing, methodology. **Adriana Notaras:** writing – review and editing, methodology. **Deepti Gurdasani:** methodology, supervision, writing – review and editing, data curation, investigation, writing – original draft. **Helen Skouteris:** conceptualization, writing – review and editing. **Darshini Ayton:** conceptualization, writing – review and editing. **Chandini Raina MacIntyre:** conceptualization, methodology, resources, funding acquisition, writing – review and editing, supervision, data curation, investigation, writing – original draft.

### Acknowledgements

E.T., R.C., D.A.H., R.D., M.K., A.N. and D.G. are supported by the Balvi Filantropic Fund. C.R.M. is funded by NHMRC Investigator Grant

2016907. D.A.H. and A.N. are funded through MRFF (grant ID 2017048). The funding sources had no role in the study design, data collection, analysis or interpretation, reporting or publication of this study. Open access publishing facilitated by University of New South Wales, as part of the Wiley - University of New South Wales agreement via the Council of Australian University Librarians.

## Ethics Statement

This study was approved by the UNSW Human Research Ethics Committee (approval number HC220737).

## Conflicts of Interest

E.T., R.C., D.A.H., R.D., M.K., A.N., H.S. and D.A. authors declare no conflicts of interest. D.G. is a member of OzSAGE. C.R.M. is on the WHO COVID-19 Vaccine Composition Technical Advisory Group and the WHO SAGE Working Group on Smallpox and Monkeypox. She receives funding from Sanofi for influenza and pertussis research and from NHMRC and MRFF.

## Data Availability Statement

The authors have nothing to report.

## References

1. S. L. Chang, Q. D. Nguyen, A. Martiniuk, V. Sintchenko, T. C. Sorrell, and M. Prokopenko, "Persistence of the Omicron Variant of SARS-CoV-2 in Australia: The Impact of Fluctuating Social Distancing. PLOS Glob," *PLOS Global Public Health* 3, no. 4 (2023): e0001427.
2. World Health Organization. "A Clinical Case Definition of Post COVID-19 Condition by a Delphi Consensus," retrieved November 6, 2022, [https://www.who.int/publications/i/item/WHO-2019-nCoV-Post\\_COVID-19\\_condition-Clinical\\_case\\_definition-2021.1](https://www.who.int/publications/i/item/WHO-2019-nCoV-Post_COVID-19_condition-Clinical_case_definition-2021.1).
3. D. Munblit, T. R. Nicholson, D. M. Needham, et al., "Studying the Post-COVID-19 Condition: Research Challenges, Strategies, and Importance of Core Outcome Set Development," *BMC Medicine* 20, no. 1 (2022): 50.
4. C. Chen, S. R. Haupt, L. Zimmermann, X. Shi, L. G. Fritsche, and B. Mukherjee, "Global Prevalence of Post-Coronavirus Disease 2019 (COVID-19) Condition or Long Covid: A Meta-Analysis and Systematic Review," *Journal of Infectious Diseases* 226, no. 9 (2022): 1593–1607.
5. A. Adjaye-Gbewonyo, A. Vahratian, C. G. Perrine, and J. Bertolli, "Long COVID in Adults: United States, 2022," NCHS Data Brief, no 480 (National Center for Health Statistics, 2023), <https://doi.org/10.15620/cdc.132417>.
6. Z. Fang, R. Ahrensbrak, and A. Rekito, "Evidence Mounts That About 7% of US Adults Have Had Long COVID," *Journal of the American Medical Association* 332, no. 1 (2024): 5–6.
7. National Center for Health Statistics. U.S. Census Bureau, Household Pulse Survey, 2022–2024. Long COVID. 2024, <https://www.cdc.gov/nchs/covid19/pulse/long-covid.htm>.
8. Office for National Statistics. Prevalence of Ongoing Symptoms Following Coronavirus (COVID-19) Infection in the UK: 30 March 2023, <https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/bulletins/prevalenceofongoingsymptomsfollowingcoronaviruscovid19infectionintheuk/30march2023>.
9. A. Vahratian, D. Adjaye-Gbewonyo, J. S. Lin, and S. Saydah, "Long COVID in Children: United States, 2022," NCHS Data Brief, no 479 (National Center for Health Statistics, 2023), <https://doi.org/10.15620/cdc.132416>.
10. Z. Al-Aly, H. Davis, L. McCorkell, et al., "Long COVID Science, Research and Policy," *Nature Medicine* 30, no. 8 (2024): 2148–2164.
11. S. R. Stein, S. C. Ramelli, A. Grazioli, et al., "SARS-CoV-2 Infection and Persistence in the Human Body and Brain at Autopsy," *Nature* 612 (2022): 758–763.
12. C. Fernández-de-las-Peñas, A. V. Raveendran, R. Giordano, and L. Arendt-Nielsen, "Long COVID or post-COVID-19 Condition: Past, Present and Future Research Directions," *Microorganisms* 11, no. 12 (2023): 2959.
13. L. L. O'Mahoney, A. Routen, C. Gillies, et al., "The Prevalence and Long-Term Health Effects of Long Covid Among Hospitalised and Non-Hospitalised Populations: A Systematic Review and Meta-Analysis," *eClinicalMedicine* 55 (2023): 101762.
14. H. E. Davis, L. McCorkell, J. M. Vogel, and E. J. Topol, "Long COVID: Major Findings, Mechanisms and Recommendations," *Nature Reviews Microbiology* 21, no. 3 (2023): 133–146.
15. C. E. Hastie, D. J. Lowe, A. McAuley, et al., "Outcomes Among Confirmed Cases and a Matched Comparison Group in the Long-COVID in Scotland Study," *Nature Communications* 13, no. 1 (2022): 5663.
16. J. Luo, J. Zhang, H. T. Tang, et al., "Prevalence and Risk Factors of Long COVID 6–12 Months After Infection With the Omicron Variant Among Nonhospitalized Patients in Hong Kong," *Journal of Medical Virology* 95, no. 6 (2023): e28862.
17. M. S. Durstenfeld, M. J. Peluso, N. D. Peyser, et al., "Factors Associated With Long COVID Symptoms in an Online Cohort Study," *Open Forum Infectious Diseases* 10, no. 2 (2023): ofad047.
18. M. Woodrow, C. Carey, N. Ziauddeen, et al., "Systematic Review of the Prevalence of Long COVID," *Open Forum Infectious Diseases* 10, no. 7 (2023): ofad233.
19. A. Subramanian, K. Nirantharakumar, S. Hughes, et al., "Symptoms and Risk Factors for Long COVID in Non-Hospitalized Adults," *Nature Medicine* 28, no. 8 (2022): 1706–1714.
20. K. Sunata, J. Miyata, H. Terai, et al., "Asthma Is a Risk Factor for General Fatigue of Long COVID in Japanese Nation-Wide Cohort Study," *Allergy International* 73, no. 2 (2024): 206–213.
21. L. Adler, S. Gazit, Y. Pinto, et al., "Long-COVID in Patients With a History of Mild or Asymptomatic SARS-CoV-2 Infection: A Nationwide Cohort Study," *Scandinavian Journal of Primary Health Care* 40, no. 3 (2022): 342–349.
22. A. Bostanci, U. Gazi, O. Tosun, et al., "Long-COVID-19 in Asymptomatic, Non-Hospitalized, and Hospitalized Populations: A Cross-Sectional Study," *Journal of Clinical Medicine* 12, no. 7 (2023): 2613.
23. K. E. Kip, E. K. McCreary, K. Collins, et al., "Evolving Real-World Effectiveness of Monoclonal Antibodies for Treatment of COVID-19: A Cohort Study," *Annals of Internal Medicine* 176, no. 4 (2023): 496–504.
24. M. Griesel, C. Wagner, A. Mikolajewska, et al., "Inhaled Corticosteroids for the Treatment of COVID-19," *Cochrane Database of Systematic Reviews* 3, no. 3 (2022): 015125.
25. C. T. R. Vegivinti, K. W. Evanson, H. Lyons, et al., "Efficacy of Antiviral Therapies for COVID-19: A Systematic Review of Randomized Controlled Trials," *BMC Infectious Diseases* 22, no. 1 (2022): 107.
26. G. Sun, K. Lin, J. Ai, and W. Zhang, "The Efficacy of Antivirals, Corticosteroids, and Monoclonal Antibodies as Acute COVID-19 Treatments in Reducing the Incidence of Long COVID: A Systematic Review and Meta-Analysis," *Clinical Microbiology and Infection* 30 (2024): 1505–1513.
27. M. Woldegiorgis, G. Cadby, S. Ngeh, et al., "Long COVID in a Highly Vaccinated but Largely Unexposed Australian Population Following the 2022 SARS-CoV-2 Omicron Wave: A Cross-Sectional Survey," *Medical Journal of Australia* 220, no. 6 (2024): 323–330.
28. P. Bager, J. Wohlfahrt, S. Bhatt, et al., "Risk of Hospitalisation Associated With Infection With SARS-CoV-2 Omicron Variant Versus Delta Variant in Denmark: An Observational Cohort Study," *Lancet Infectious Diseases* 22, no. 7 (2022): 967–976.
29. M. Antonelli, J. C. Pujol, T. D. Spector, S. Ourselin, and C. J. Steves, "Risk of Long COVID Associated With Delta Versus Omicron Variants of SARS-CoV-2," *Lancet* 399, no. 10343 (2022): 2263–2264.

30. S. M. Pinto Pereira, A. Mensah, M. D. Nugawela, et al., “Long COVID in Children and Young After Infection or Reinfection With the Omicron Variant: A Prospective Observational Study,” *Journal of Pediatrics* 259 (2023): 113463.
31. A. Adekunle, M. Meehan, D. Rojas-Alvarez, J. Trauer, and E. McBryde, “Delaying the COVID-19 Epidemic in Australia: Evaluating the Effectiveness of International Travel Bans,” *Australian and New Zealand Journal of Public Health* 44, no. 4 (2020): 257–259.
32. Australian Government Department of Health and Aged Care. COVID-19 Vaccination Rollout Update. 2021, <https://www.health.gov.au/resources/collections/covid-19-vaccination-rollout-update#december-2022>.
33. Dynata. Dynata: World’s Largest First Party Data Platform 2023, <https://www.dynata.com/>.
34. R. C. S. Danielle, D. M. Débora, N. L. P. Alessandra, et al., “Correlating COVID-19 Severity With Biomarker Profiles and Patient Prognosis,” *Scientific Reports* 14, no. 1 (2024): 22353.
35. A. L. Cabrera Martimbianco, R. L. Pacheco, A. M. Bagattini, and R. Riera, “Frequency, Signs and Symptoms, and Criteria Adopted for Long COVID-19: A Systematic Review,” *International Journal of Clinical Practice* 75, no. 10 (2021): e14357.
36. O. Byambasuren, P. Stehlik, J. Clark, K. Alcorn, and P. Glasziou, “Effect of Covid-19 Vaccination on Long Covid: Systematic Review,” *BMJ Medicine* 2, no. 1 (2023): e000385.
37. Y. Xie, B. Bowe, and Z. Al-Aly, “Molnupiravir and Risk of Hospital Admission or Death in Adults With Covid-19: Emulation of a Randomized Target Trial Using Electronic Health Records,” *BMJ* 380 (2023): e072705.
38. T. C. Lee, É. Bortolussi-Courval, S. Belga, et al., “Inhaled Corticosteroids for Outpatients With COVID-19: A Meta-Analysis,” *European Respiratory Journal* 59, no. 5 (2022): 2102921.
39. Australian Institute of Health and Welfare, *Long COVID in Australia—a Review of the Literature, Catalogue Number PHE 318* (AIHW, Australian Government, 2022).
40. Y. Xie, B. Bowe, and Z. Al-Aly, “Burdens of Post-Acute Sequelae of COVID-19 by Severity of Acute Infection, Demographics and Health Status,” *Nature Communications* 12, no. 1 (2021): 6571.
41. K. Mølbak, T. I. A. Sørensen, S. Bhatt, F. P. Lyngse, L. Simonsen, and P. Aaby, “Severity of Respiratory Tract Infections Depends on the Infectious Dose. Perspectives for the Next Pandemic,” *Front Public Health* 12 (2024): 1391719.
42. J. C. Stoltzfus, “Logistic Regression: A Brief Primer,” *Academic Emergency Medicine* 18, no. 10 (2011): 1099–1104.
43. Y. Benjamini and Y. Hochberg, “Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing,” *Journal of the Royal Statistical Society Series B: Statistical Methodology* 57, no. 1 (1995): 289–300.
44. T. J. Loftus, B. Shickel, J. A. Balch, et al., “Phenotype Clustering in Health Care: A Narrative Review for Clinicians,” *Frontiers in Artificial Intelligence* 5 (2022): 842306.
45. R Core Team, *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, 2024).
46. M. Maechler, P. Rousseeuw, A. Struyf, M. Hubert, and K. Hornik, *Cluster Analysis Basics and Extensions*. R Package Version 2.1.6. 2023.
47. Hamming Distance, *Encyclopedia of Biometrics*, ed. S. Z. Li and A. Jain (Springer US, 2009), 668.
48. P. J. Rousseeuw, “Silhouettes: A Graphical Aid to the Interpretation and Validation of Cluster Analysis,” *Journal of Computational and Applied Mathematics* 20 (1987): 53–65.
49. The Mathworks Inc. MATLAB version: 9.14.0.2337262 (R2023a). 2022, <https://www.mathworks.com>.
50. T. Bahmer, C. Borzikowsky, W. Lieb, et al., “Severity, Predictors and Clinical Correlates of Post-COVID Syndrome (PCS) in Germany: A Prospective, Multi-Centre, Population-Based Cohort Study,” *eClinicalMedicine* 51 (2022): 101549.
51. M. Whitaker, J. Elliott, M. Chadeau-Hyam, et al., “Persistent COVID-19 Symptoms in a Community Study of 606,434 People in England,” *Nature Communications* 13, no. 1 (2022): 1957.
52. S. L. Klein and K. L. Flanagan, “Sex Differences in Immune Responses,” *Nature Reviews Immunology* 16, no. 10 (2016): 626–638.
53. O. Magen, J. G. Waxman, M. Makov-Assif, et al., “Fourth Dose of BNT162b2 mRNA Covid-19 Vaccine in a Nationwide Setting,” *New England Journal of Medicine* 386, no. 17 (2022): 1603–1614.
54. Australian Government Department of Health and Aged Care. “ATAGI Update on the COVID-19 Vaccination Program,” retrieved October 6, 2023, <https://www.health.gov.au/news/atagi-update-on-the-covid-19-vaccination-program?language=en>.
55. Y. Xie, T. Choi, and Z. Al-Aly, “Nirmatrelvir and the Risk of Post-Acute Sequelae of COVID-19,” preprint, *MedRxiv*, 2022, <https://doi.org/10.1101/2022.11.03.22281783>.
56. Y. Xie, T. Choi, and Z. Al-Aly, “Molnupiravir and Risk of Post-Acute Sequelae of Covid-19: Cohort Study,” *BMJ* 381 (2023): 074572.
57. G. Kenny, K. McCann, C. O’Brien, et al., “Identification of Distinct Long Covid Clinical Phenotypes Through Cluster Analysis of Self-Reported Symptoms,” *Open Forum Infectious Diseases* 9, no. 4 (2022): ofac060, <https://doi.org/10.1093/ofid/ofac060>.
58. Australian Bureau of Statistics. National, State and Territory Population. 2023, <https://www.abs.gov.au/statistics/people/population/national-state-and-territory-population/latest-release#states-and-territories>.
59. Australian Bureau of Statistics. A Caring Nation—15 Per Cent of Australia’s Workforce in Health Care and Social Assistance Industry. 2022, <https://www.abs.gov.au/media-centre/media-releases/caring-nation-15-cent-australias-workforce-health-care-and-social-assistance-industry>.
60. Z. Akhtar, M. Trent, A. Moa, T. C. Tan, O. Fröbert, and C. R. MacIntyre, “The Impact of COVID-19 and COVID Vaccination on Cardiovascular Outcomes,” supplement, *European Heart Journal Supplements* 25, no. Suppl A (2023): A42–A49.

## Supporting Information

Additional supporting information can be found online in the Supporting Information section.