

Individual, household structure, and socioeconomic predictors of COVID-19 testing and vaccination outcomes: a whole population linked data analysis

Nicole Satherley^{1,2*} and Andrew Sporle^{1,2}

Submission History

Submitted:	23/10/2024
Accepted:	28/03/2025
Published:	27/05/2025

¹ iNZight Analytics Ltd

² Department of Statistics,
University of Auckland

Abstract

Introduction

The COVID-19 pandemic produced social inequities in health outcomes between and within nations. Reported inequitable COVID-19 outcomes for ethnic minorities and indigenous peoples are likely to be associated in part because of poorer socioeconomic circumstances experienced by these populations. Understanding these associations within national populations is vital for future pandemic management.

Objective

This study explores the social inequity of COVID-19 outcomes within New Zealand over the first 3 years of the pandemic. We aimed to identify policy amenable socioeconomic factors associated with COVID-19 outcomes while adjusting for relevant individual factors and household structure. We also aimed to examine whether ethnic group differences are smaller when accounting for these socioeconomic factors and household structure.

Methods

Administrative individual-level data for the New Zealand population was analysed to assess COVID-19 health outcomes during 2020 – 2023. The association between individual (e.g. age, ethnicity, disability status), household structure (e.g. household composition) and socioeconomic (e.g. crowding, housing quality, deprivation) factors and four COVID-19 health outcomes – infection, hospitalisation, mortality, and vaccination status was assessed.

Results

Indigenous peoples and ethnic minorities experienced worse outcomes across most COVID-19 outcomes. Adjusting for household structure and socioeconomic factors reduced but did not eliminate these inequities between ethnic groups. Housing issues including high housing mobility, poor quality housing, and household crowding were associated with worse outcomes, as were disability status, no primary health care enrolment, lower household income and older age. The size of these effects also differed for different health outcomes.

Conclusions

Ethnic inequity was persistent and likely partly explained by policy-modifiable social factors, despite the relatively minor population health impacts of COVID-19 in New Zealand. We also demonstrate how a range of socioeconomic determinants predict COVID-19 outcomes in different ways.

Keywords

COVID-19; New Zealand; administrative data; social inequity; social determinants

*Corresponding Author:

Email Address: nicole@inzight.co.nz (Nicole Satherley)

Highlights

- Whole-population administrative data analysis of a wide range of Covid-19 outcomes and determinants in New Zealand.
- Ethnic group inequity was identified despite the relatively limited population impact of Covid-19 in New Zealand.
- Policy modifiable socioeconomic factors are widely predictive of Covid-19 health outcomes.
- Accounting for socioeconomic but not household structure factors reduced ethnic group differences in health outcomes.

Introduction

Understanding who in a population is most likely to contract COVID-19 and experience poorer health outcomes is important for managing pandemic response and minimising impact. Studies into individual risk factors have consistently identified factors such as older age and health comorbidities as increasing the risk of severe COVID-19 infection and mortality [1–3]. The potential for inequitable *ethnic* outcomes during the COVID-19 pandemic were also clearly identified early on. Despite early policy recommendations to protect indigenous populations [4], ethnic inequities have been widely reported during the pandemic. Inequitable outcomes for indigenous populations have also been identified, but they have been more difficult to and rarely assessed due to a lack of available high-quality data [5, 6]. Ethnic minorities within and between nations are generally more likely to be infected with COVID-19, experience greater severity of infection, and higher mortality rates compared to their ethnic majority counterparts [7–12].

Even seemingly successful national COVID-19 pandemic strategies have not addressed and prevented ethnic inequities in outcomes. New Zealand's COVID-19 response was accompanied by *negative* excess mortality across the whole population during 2020–2021 [13]. Yet ethnic inequities in COVID-19 outcomes occurred despite available and longstanding evidence of ethnic inequity in morbidity [14], ethnic inequity during prior pandemics [15, 16], and early warnings during the COVID-19 pandemic that Māori and Pacific peoples would be at increased risk [17, 18]. Initial local transmission of COVID-19 in New Zealand mostly involved Pacific populations [19]. Early pandemic data confirmed Māori and Pacific peoples had a higher infection rate and a higher hospitalisation rate relative to other New Zealanders [20, 21]. Despite this, Māori and Pacific peoples tended to have poorer spatial access to vaccination centres [22]. This inequity has since been the subject of a government commission of inquiry which identified several failings by the government COVID-19 response to protect Māori [23].

Individual factors in COVID-19 outcomes such as ethnicity are important to assess as they can inform interventions to minimise impact and inequity. However, socioeconomic circumstances such as deprivation, crowded households and poor-quality housing are also highly relevant to health

outcomes [24] and help to explain ethnic group differences in COVID-19 outcomes [12, 13, 25–27]. Geographic areas with lower socioeconomic status scores in Santiago, Chile, experienced higher death rates [28], while areas with poorer housing conditions and overcrowding had higher COVID-19 incidence and mortality [29–31]. Household structure may also play a role in increased household transmission [24, 32]. For example, living in intergenerational households (such as elderly living with children) has been associated with more suspected COVID-19 cases within an area [31] and a higher risk of mortality among the elderly [33]. Transmission is greater in larger households [34], and severe COVID-19 infection more common in both larger and single person households [35].

Incorporating socioeconomic and household structure factors into COVID-19 outcome models is also important because unlike many individual factors, factors such as housing quality and crowding are policy amenable. However, comprehensive, large-scale analysis of COVID-19 health outcomes among populations is challenging because it requires high quality data at the individual level on both COVID-19 outcomes (e.g. cases, hospitalisations, deaths, vaccinations), and relevant demographic and socioeconomic data that can be linked to the people experiencing these outcomes. In this study we overcome these challenges and examine multiple COVID-19 health outcomes in New Zealand using whole population administrative data.

The New Zealand context provides an ideal environment for studying COVID-19 outcomes due to the combination of our isolation, single level of government, service provision, and high quality, whole nation linked administrative data. The closure of the border to non-citizens and residents for most of the early pandemic (to July 2022) means transmission of and outcomes for COVID-19 were largely due to local conditions and decision making (rather than policy decisions by neighbouring nations). There was a single uniform pandemic strategy for the entire nation, (although different regions were in lockdown at different points in time). Universal access to public healthcare makes it more likely that people with more severe infection would present to hospitals with admission and outcomes recorded in official statistics.

High quality official data is available on COVID-19 test results (supervised PCR tests), vaccination records, and hospitalisations and deaths, with mandatory recording of hospital admissions and registrations of death. This data is available alongside linked de-identified administrative microdata in Stats NZ's Integrated Data Infrastructure (IDI), [36]. This study draws on a wide range of administrative data sources in the IDI to assess factors relevant to COVID-19 infection and health outcomes at the individual-level for the entire New Zealand population. Our assessment of multiple outcomes and relevant factors in unified models within a consistent population is a novel addition to previous studies that focused on limited factors at once, area-level associations, or non-administrative data sources.

This study examines the association between four COVID-19 outcomes (testing positive, hospitalisation, death, and vaccination status) and individual factors (Māori and Pacific ethnicity, age, sex, disability status, and enrolment with a primary healthcare provider), household structure (intergenerational living, household composition, and the presence of a child under the age of 5), and policy amenable

socioeconomic factors (area level deprivation index score, household income, household crowding, housing quality, and housing mobility) in New Zealand using the IDI. We aimed to identify policy amenable socioeconomic factors associated with COVID-19 outcomes while adjusting for individual factors and household structure. We also aimed to observe whether ethnic group differences in COVID-19 outcomes are smaller when accounting for these policy amenable socioeconomic factors and household structure.

Method

Data and study population

This work uses a collection of de-identified administrative microdata from various government agencies in the IDI. Data was sourced from the 2018 Census, Ministry of Health (MOH), and core data derived by Stats NZ, including address notifications. Figure 1 demonstrates the data joining process to obtain the variables for the analyses. Because the 2018 Census provided most of the data for predictor variables, we used the 2018 Census Usually Resident Population (URP) as our population of interest. This population was determined by the 2018 Census of Population and Dwellings held on 6 March 2018, and consists of 4,699,755 individuals who usually live in New Zealand and were present in the country on Census night, excluding visitors from overseas and residents temporarily overseas. The 2018 Census counted 97.4% of people usually resident in New Zealand on census night based on a combination of individual form completions, household listings, and administrative data enumerations. More information on the 2018 Census and

demographic details of the 2018 URP can be found at <https://www.stats.govt.nz/2018-census/>.

Not everyone in the 2018 URP had the required data to be included in our analyses. Specifically, among the 2018 URP, 1,851,537 individuals had COVID-19 testing data (i.e. had taken at least one COVID-19 test) by 16 February 2022, of which 1,333,944 had complete data for the analysis. For assessing hospitalisations, 983,709 individuals had tested positive for COVID-19 by 2 June 2022 of which 697,614 had complete data on independent variables and had not died within one month of a positive test without a record of hospitalisation (567 people). For assessing deaths, 1,754,076 had tested positive by 14 February 2023, of which 1,314,810 had complete data on independent variables. In addition, after excluding those who had died (111,297) or left New Zealand without return (161,211) by 10 May 2023 and were not fully vaccinated, 4,427,844 people had a vaccination status. Demographic details of each analysis sample, and corresponding details in the total 2018 Census URP are displayed in Table 1. Sample and total URP demographics were comparable, but our samples had a slightly higher proportion of women and, depending on the sample, a lower proportion of Māori and Pacific peoples.

Measures

COVID-19 Outcome variables

COVID-19 test result

The COVID-19 test result variable included all COVID-19 tests from 30 April 2020 up to 16 February 2022 in New Zealand

Figure 1: Summary of data joining steps for variables in the Integrated Data Infrastructure (IDI)

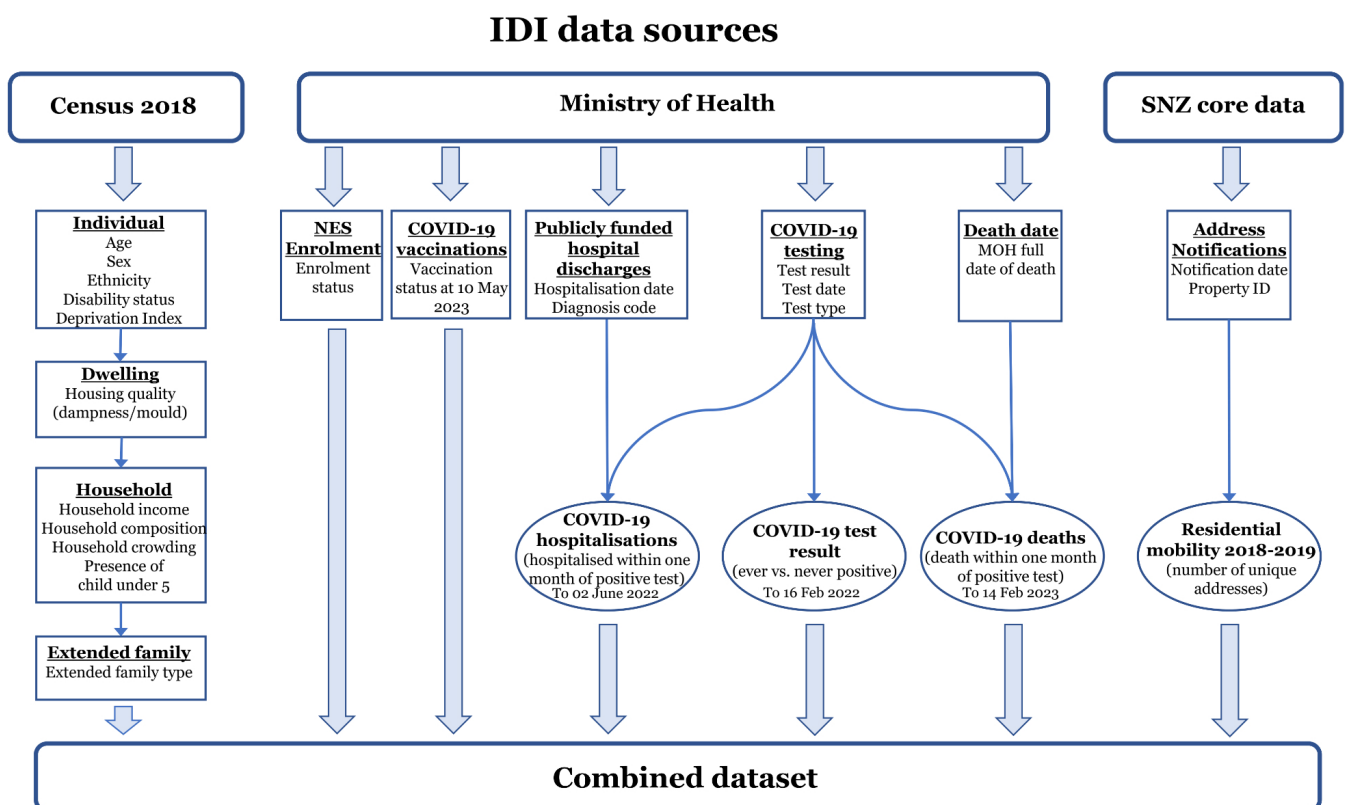


Table 1: Final analysis sample sizes and demographics compared to the full 2018 Census Usually Resident Population

	Test result	Hospitalisation	Death	Full vs. no vaccination	Full vs. partial vaccination	Partial vs. no vaccination	2018 Census URP
Analysis sample size	1,333,944	697,614	1,314,810	3,060,945	2,810,535	359,766	–
Age (mean)	39.3	32.6	37.3	41.4	41.6	32.1	37.4
Women (%)	54.6%	54.9%	55.2%	52.1%	51.9%	52.1%	50.6%
Māori (%)	14.1%	16.5%	13.4%	12.9%	12.3%	21.2%	16.5%
Pacific (%)	7.7%	8.9%	6.2%	5.5%	5.4%	8.1%	8.1%

Note. URP = Usually Resident Population. For information on the 2018 Census URP see <https://www.stats.govt.nz/tools/2018-census-place-summaries/new-zealand>.

MOH records and was recorded as whether someone had never (0) or ever (1) recorded a positive COVID-19 test. The cut-off date was used as unsupervised self-testing commenced after this (test results may not be officially recorded beyond this point).

COVID-19 hospitalisation

COVID-19 hospitalisations were defined as whether someone had never (0) or ever (1) been hospitalised for COVID-19 (with any COVID-19 related ICD-10-AM code) within one month of a positive COVID-19 test, for positive tests up to 02 June 2022. This cut off point was required as hospitalisation data supplied by the MOH is only available up to 30 June 2022 in the March 2023 IDI refresh (thus, assessing hospitalisation outcomes for tests beyond this point was not possible).

COVID-19 death

COVID-19 deaths were defined as whether someone did not (0) or did (1) die within one month of a reported positive COVID-19 test result, for test results up to 14 February 2023. The February 2023 cut off point was chosen and required as MOH death records were available up to March 2023. Officially confirmed cause of death was unavailable in the IDI beyond 2018, so these deaths represent attributed, but not confirmed, COVID-19 related deaths. This coding approach is consistent with the MOH [37].

COVID-19 vaccination status

Vaccination status used codes provided directly from the MOH Covid Immunisation Register for whether someone was fully vaccinated (1) or partially vaccinated (2) up to 10 May 2023, the latest date of available records in the IDI. Individuals who did not have a vaccination record in the MOH data (and who had not left the country or died by 10 May 2023) were assumed to be unvaccinated (0). Vaccination eligibility varied by time and age group during New Zealand's initial vaccination program. Eligibility for vaccine doses (e.g. based on age) is accounted for within MOH vaccination status records [38].

Household structure variables

Household composition was obtained through the 2018 Census and coded into the following categories: one-family household, two or more family household, other multi-person household, and one-person household.

Extended family types were obtained from the 2018 Census, and coded as either living with three or more generations of extended family (e.g. couple with children and grandchildren, or grandparents and grandchildren; coded 1) or not (coded 0).

The presence of a child under 5 was based on the age of the youngest child in a family/household in the 2018 Census. A code of 0 indicates no child under 5 was present, and a code of 1 indicates a child under 5 was present. Note that because COVID-19 outcomes occurred during 2020 – 2023, children will be under the age of 10 in a household at the time the COVID-19 outcome occurred. As such, the variable captures a 0 – 5-year-old cohort at the time of the 2018 Census, but will miss children born after the 2018 Census, and may therefore underestimate the effect of the presence of a child under 5.

Socioeconomic variables

Household crowding was coded from the 2018 Census variable used by Stats NZ, based on the Canadian National Occupancy Standard [39]. Those not needing additional bedrooms were coded 0 (not crowded), and those needing one or more additional bedrooms coded 1 (crowded).

Housing quality was derived from the dwelling dampness and mould variables in the 2018 Census. A code of 1 represents a dwelling always damp or that always had mould over A4 in size, and a code of 0 represents the presence of dampness or mould sometimes or never.

Jensen Equivalised household income was obtained from the 2018 Census but converted to \$10,000 units for ease of interpreting regression coefficients (each unit increase in income corresponds to \$10,000).

Residential mobility was derived as a count of the number of unique property IDs recorded for each individual during 2018 – 2019 (prior to the pandemic) in the address notification table. This data table contains a list of residential addresses reported for an individual obtained through their interaction with various government agencies, collated by Stats NZ. Individuals without a record in the address notification table during this period were coded as 0. Residential mobility was then categorised as no moves or one move (0; low mobility), two or three moves (1; medium mobility) or 4 or more moves (2; high mobility) during the two-year period.

The 2018 New Zealand deprivation index was used from the 2018 Census, and ranges from 1 (least deprived) to 10 (most deprived). This is a small area-level measure of deprivation based on characteristics such as the proportion

of people in that area with no home internet access, people aged 18 – 64 without any qualifications, and people not living in their own home [40].

Individual predictor variables

Age in years, birth month and year, and sex (0 = men, 1 = women) were obtained from the 2018 Census. Age in years at the 2018 Census was used for the regression models but was divided by 10 to better convey the effect size for age in the results. Thus, each unit increase corresponds to each 10-year increase in age.

Māori ethnicity was coded for those recorded as Māori only or Māori and at least one other ethnic group in the 2018 Census (scored 1, non-Māori = 0). Pacific ethnicity was similarly coded 1 for those who were recorded as sole Pacific or Pacific and at least one other ethnic group (non-Pacific = 0). The comparison group for the ethnicity variables are non-Māori and non-Pacific peoples respectively (predominantly NZ European).

Disability status (0 = no, 1 = yes) was from 2018 Census responses to the Washington Group Short Set, and coded as having a lot of difficulty or being unable to do one or more of six activities (seeing, hearing, walking or climbing steps, remembering or concentrating, washing all over or dressing, communicating).

Primary health organisation (PHO) enrolment was derived based on National Enrolment Service (NES) records as an indicator variable (not enrolled = 0, enrolled = 1) based on any active enrolment during 2020 (for COVID-19 testing outcome, hospitalisation, and death regressions), or based on enrolment status in the month prior (July 2021) to the start of the vaccination rollout for the general population.

Analytic approach

Logistic regression models were used to analyse positive COVID-19 tests (ever testing positive vs. never testing positive), COVID-19 hospitalisations (being hospitalised within one-month of a positive test vs. not being hospitalised), COVID-19 deaths (death within one-month of a positive test vs. no death), and vaccination status (being fully vaccinated vs. not vaccinated, partially vaccinated vs. not vaccinated, and partially vaccinated vs. fully vaccinated). Because the data consists of people within households experiencing similar circumstances, we used Generalised Estimating Equations [41] with household as the cluster variable to account for potential non-independence in model residuals. The main full models are presented first, and then compared to models with only individual, and only individual and household structure variables entered. This allowed for an initial descriptive assessment of whether household structure and socioeconomic factors might partially explain (reduce) ethnic group differences in outcomes.

Results

Unadjusted risk of COVID-19 outcomes

Of the 1,851,537 people in the 2018 Census URP who had ever recorded a COVID-19 test result up to 16 February 2022,

16,896 (1%) tested positive. Of the 983,709 who had ever tested positive, up to 02 June 2022, 10,404 (1.1%) had been subsequently hospitalised within one month. Finally, of the 1,754,076 who had ever tested positive up to 14 February 2023, 3,489 (0.2%) had subsequently died within one month. The overall risk of hospitalisation or death within one month of a positive COVID-19 test was therefore low, as was the positive test rate for the first two years of the pandemic.

In terms of vaccinations, when excluding those who migrated or died before 10 May 2023, 81.4% (3,603,495) were fully vaccinated for COVID-19, 3.0% (133,434) were partially vaccinated, and 15.6% (690,918) were unvaccinated, as of 10 May 2023.

COVID-19 infection outcomes

Demographic and socioeconomic factors were widely predictive of COVID-19 infection outcomes (Table 2). Māori had 1.17 times (95% CI = 1.08–1.26) the odds of non-Māori, and Pacific peoples 1.93 times (95% CI = 1.76–2.12) the odds of non-Pacific people of testing positive. People with a disability had lower odds of testing positive (OR = 0.85, 95% CI = 0.74–0.97).

Living with a child under 5 (OR = 1.16, 95% CI = 1.08–1.25) or with three or more generations of extended family (OR = 1.18, 95% CI = 1.04–1.34) were associated with greater odds of testing positive. Those living alone had lower odds of testing positive than those living in a single-family household (OR = 0.76, 95% CI = 0.66–0.86). Those living in a more deprived area had higher odds (OR = 1.10, 95% CI = 1.08–1.11) and those with a higher household income had lower odds (OR = 0.99, 95% CI = 0.98–1.00) of testing positive. Those with high residential mobility, relative to low mobility, had 1.47 times greater odds of testing positive (95% CI = 1.27–1.71), while those living in crowded households had 1.59 times greater odds (95% CI = 1.45–1.74) and those living in poorer quality housing 1.19 times greater odds (95% CI = 1.08–1.32).

Comparison regression models show the ethnic group differences were stronger when only individual factors were entered into the models (for Māori, OR = 1.28; for Pacific peoples, OR = 2.21; Supplementary Table 1) but these were unchanged when including household structure variables (Supplementary Table 2). This indicates that it was the inclusion of the socioeconomic variables specifically that reduced the point-estimate of ethnic group differences in odds of testing positive.

For hospitalisations (centre columns of Table 2), Māori had 1.19 (95% CI = 1.11–1.27) and Pacific peoples 1.55 (95% CI = 1.41–1.66) times the odds of non-Māori and non-Pacific peoples respectively of being hospitalised following a positive COVID-19 test. Those enrolled with a PHO had higher odds of being hospitalised (OR = 1.72, 95% CI = 1.38–2.13) and those with a disability had 2.37 times greater odds of being hospitalised (95% CI = 2.20–2.55). Those living alone had higher odds of being hospitalised than those living with a single family (OR = 1.30, 95% CI = 1.20–1.41).

In contrast to household structure, socioeconomic factors were widely associated with hospitalisation status. Those in a more deprived area had greater odds (OR = 1.04, 95% CI = 1.03–1.05), and those with greater household

Table 2: Individual, household structure, and socioeconomic predictors of COVID-19 infection outcomes

	Positive covid test (vs. negative covid test) for tests to 16 February 2022				Hospitalisation following positive test (vs. not hospitalised) for positive tests to 02 June 2022				Death following positive test (vs. no death) for positive tests to 14 February 2023			
	Odds ratio	95% CI Lower	95% CI Upper	<i>p</i>	Odds ratio	95% CI Lower	95% CI Upper	<i>p</i>	Odds ratio	95% CI Lower	95% CI Upper	<i>p</i>
Māori Ethnicity (0 no, 1 yes)	1.17	1.08	1.26	<.001	1.19	1.11	1.27	<.001	1.31	1.12	1.52	.001
Pacific Ethnicity (0 no, 1 yes)	1.93	1.76	2.12	<.001	1.53	1.41	1.66	<.001	1.25	0.99	1.57	.059
Sex (0 women, 1 men)	1.01	0.96	1.06	.717	0.82	0.78	0.86	<.001	1.77	1.63	1.93	<.001
Age	0.91	0.89	0.92	<.001	1.49	1.46	1.51	<.001	3.25	3.11	3.39	<.001
PHO enrolment (0 no, 1 yes)	0.77	0.66	0.89	.001	1.72	1.38	2.13	<.001	0.58	0.38	0.89	.012
Disability status (0 no, 1 yes)	0.85	0.74	0.97	.020	2.37	2.20	2.55	<.001	2.15	1.95	2.38	<.001
Child under 5 (0 no, 1 yes)	1.16	1.08	1.25	<.001	0.90	0.84	0.97	.003	0.91	0.71	1.18	.487
Extended family (0 no, 1 yes)	1.18	1.04	1.34	.009	1.15	1.03	1.28	.010	1.32	1.06	1.63	.012
<i>Household composition^a</i>												
Two or more families	1.13	1.00	1.28	.048	1.08	0.96	1.20	.188	1.08	0.85	1.38	.504
Multi-person	0.90	0.79	1.02	.111	1.11	0.98	1.26	.116	1.29	1.01	1.65	.042
One person	0.76	0.66	0.86	<.001	1.30	1.20	1.41	<.001	1.24	1.12	1.38	<.001
Deprivation	1.10	1.08	1.11	<.001	1.04	1.03	1.05	<.001	1.04	1.02	1.06	<.001
Household income	0.99	0.98	1.00	.001	0.95	0.94	0.95	<.001	0.94	0.93	0.95	<.001
<i>Residential mobility^b</i>												
Medium	1.10	1.02	1.17	.010	1.17	1.10	1.24	<.001	1.05	0.94	1.16	.395
High	1.47	1.27	1.71	<.001	1.91	1.68	2.16	<.001	1.68	1.30	2.18	<.001
Household crowding (0 no, 1 yes)	1.59	1.45	1.74	<.001	1.20	1.10	1.31	<.001	1.10	0.86	1.39	.447
Housing quality (0 never or sometimes damp or mould, 1 always)	1.19	1.08	1.32	<.001	1.25	1.15	1.37	<.001	1.10	0.87	1.38	.440

Note. $N(\text{test outcome}) = 1,333,944$, $N(\text{hospitalisation outcome}) = 697,614$. $N(\text{death outcome}) = 1,314,810$. * $p < .05$, ** $p < .01$, *** $p < .001$. ^aReference category is those in single family households. ^bReference category is those with low mobility.

incomes had lower odds (OR = 0.95, 95% CI = 0.94–0.95) of hospitalisation. Individuals with high residential mobility had 1.91 times greater odds of being hospitalised (95% CI = 1.68–2.16) than those with low mobility. Finally, both household crowding (OR = 1.20, 95% CI = 1.10–1.31) and poorer housing quality (OR = 1.25, 95% CI = 1.15–1.37) were each associated with greater odds of being hospitalised.

Alternative models that included individual factors alone indicated larger ethnic group differences in odds of hospitalisation (Māori OR = 1.43, Pacific OR = 1.98; Supplementary Table 1), compared to the full model. These effects did not reduce with the inclusion of household structure variables alone (Supplementary Table 2). This indicates that it was the inclusion of socioeconomic variables specifically that reduced the point-estimate of ethnic group differences in odds of hospitalisation. This could indicate that the socioeconomic variables examined here could better explain the ethnic group differences than the household structure variables examined.

Finally, model results for death following a positive COVID-19 test are displayed on the right-hand side of Table 2. Māori had 1.31 times (95% CI = 1.12–1.52) the odds of non-Māori of dying following a positive COVID-19 test. Those with a disability had 2.15 times greater odds of death (95% CI = 1.95–2.38) while those enrolled with a PHO had lower odds (OR = 0.58, 95% CI = 0.38–0.89). Men had 1.77 times greater odds of death than women (95% CI = 1.63–1.93). In terms of household structure variables, those living in a one-person household had higher odds of dying than those in a single-family household (OR = 1.24, 95% CI = 1.12–1.38). Those living in a multi-person household

(OR = 1.29, 95% CI = 1.01–1.65) and those living with three or more generations of family (OR = 1.32, 95% CI = 1.06–1.63) also had higher odds of dying, although there was more uncertainty around these estimates.

Higher area level deprivation (OR = 1.04, 95% CI = 1.02–1.06) was associated with greater odds, and higher household income (OR = 0.94, 95% CI = 0.93–0.95) was associated with lower odds of dying following a positive COVID-19 test. Those highly mobile (OR = 1.68, 95% CI = 1.30–2.18) had higher odds of dying relative to those low in residential mobility, while the remaining socioeconomic variables were generally unassociated with COVID-19 mortality.

The comparison model with only individual factors entered again shows larger odds ratio effect sizes for Māori (OR = 1.54) and Pacific (OR = 1.58) ethnicity (Supplementary Table 1). There was little reduction in the effect sizes with the addition of household structure variables in the model (Supplementary Table 2), but a noticeable reduction with the addition of socioeconomic variables. This again might suggest that the policy amenable socioeconomic factors included here could partially explain ethnic group differences in COVID-19 infection outcomes.

COVID-19 vaccination status

Results of the regression models for different COVID-19 vaccination status are presented in Table 3, with reduced models including only individual, or individual and household structure variables, presented in the Supplementary Tables 3, 4. Māori had lower odds than non-Māori of being fully vs. partially vaccinated (OR = 0.77, 95% CI = 0.75–0.78), fully vs.

Table 3: Individual, household structure, and socioeconomic predictors of vaccination status as at 10 May 2023

	Fully vaccinated (vs. Partially vaccinated)				Fully vaccinated (vs. Not vaccinated)				Partially vaccinated (vs. Not vaccinated)			
	Odds ratio	95% CI Lower	95% CI Upper	P	Odds ratio	95% CI Lower	95% CI Upper	P	Odds ratio	95% CI Lower	95% CI Upper	P
Māori Ethnicity (0 no, 1 yes)	0.77	0.75	0.78	<.001	0.76	0.75	0.77	<.001	0.90	0.88	0.93	<.001
Pacific Ethnicity (0 no, 1 yes)	1.02	0.99	1.06	.164	0.96	0.94	0.98	<.001	1.07	1.03	1.11	.001
Sex (0 women, 1 men)	0.90	0.89	0.92	<.001	1.19	1.18	1.20	<.001	1.06	1.04	1.08	<.001
Age	2.77	2.73	2.81	<.001	1.13	1.13	1.14	<.001	0.64	0.64	0.65	<.001
PHO enrolment (0 no, 1 yes)	0.97	0.93	1.01	.118	11.63	11.50	11.77	<.001	6.85	6.59	7.11	<.001
Disability status (0 no, 1 yes)	0.76	0.72	0.79	<.001	0.74	0.73	0.75	<.001	1.09	1.04	1.15	<.001
Child under 5 (0 no, 1 yes)	0.43	0.42	0.44	<.001	0.66	0.65	0.67	<.001	1.04	1.02	1.07	<.001
Extended family (0 no, 1 yes)	0.93	0.90	0.97	.001	0.86	0.84	0.88	<.001	1.14	1.09	1.19	<.001
<i>Household composition^a</i>												
Two or more families	1.07	1.02	1.12	.003	1.10	1.07	1.13	<.001	0.95	0.90	1.00	.032
Multi-person	2.00	1.85	2.17	<.001	1.40	1.36	1.43	<.001	0.63	0.58	0.68	<.001
One person	0.35	0.33	0.37	<.001	0.87	0.85	0.88	<.001	1.04	0.97	1.10	.259
Deprivation	0.98	0.98	0.98	<.001	0.99	0.98	0.99	<.001	0.99	0.98	0.99	<.001
Household income	1.04	1.04	1.04	<.001	1.05	1.05	1.05	<.001	1.02	1.02	1.02	<.001
<i>Residential mobility^b</i>												
Medium	1.12	1.09	1.14	<.001	1.22	1.21	1.23	<.001	0.91	0.89	0.93	<.001
High	1.06	1.00	1.12	.038	1.12	1.08	1.15	<.001	0.85	0.80	0.90	<.001
Household crowding (0 no, 1 yes)	1.03	1.00	1.06	.043	1.06	1.03	1.08	<.001	0.98	0.95	1.02	.310
Housing quality (0 never or sometimes damp or mould, 1 always)	0.97	0.93	1.00	.052	0.86	0.85	0.88	<.001	0.94	0.90	0.98	.002

Note. $N(\text{Fully vs. partially}) = 2,810,535$. $N(\text{Fully vs. none}) = 3,060,945$. $N(\text{Partially vs. none}) = 359,766$. * $p < .05$, ** $p < .01$, *** $p < .001$.^aReference category is those in single family households.^bReference category is those with low mobility.

not vaccinated (OR = 0.76, 95% CI = 0.75–0.77), and partially vs. not vaccinated (OR = 0.90, 95% CI = 0.88–0.93). The pattern in inequity was different for Pacific peoples, who did not differ from non-Pacific people in being fully vs. partially vaccinated (OR = 1.02, $p = .164$), and only slightly lower odds of being fully vaccinated vs. not vaccinated (OR = 0.96, 95% CI = 0.94–0.98). They also had slightly higher odds than non-Pacific peoples of being partially vs. not vaccinated (OR = 1.07, 95% CI = 1.03–1.11). Thus, by May 2023, Pacific peoples, relative to non-Pacific, generally had higher odds of having started the vaccination process but slightly lower odds of having finished the vaccination process (i.e. being fully vaccinated).

Age (each additional 10-years of age) was particularly strongly associated with greater odds of being fully vs. partially vaccinated (OR = 2.77, 95% CI = 2.73–2.81), but lower odds of being partially vs. not vaccinated (OR = 0.64, 95% CI = 0.64–0.65). There was a strong effect of PHO enrolment, such that those enrolled with a PHO had 11.63 times greater odds of being fully vs. not vaccinated (95% CI = 11.50–11.77) and 6.85 times greater odds of being partially vs. not vaccinated (95% CI = 6.59–7.11), but enrolment status was unassociated with being fully vs. partially vaccinated (OR = 0.97, 95% CI = 0.93–1.01).

Living with a child under the age of 5 was associated with lower odds of being fully vs. partially (OR = 0.43, 95% CI = 0.42–0.44) or not vaccinated (OR = 0.66, 95% CI = 0.65–0.67), and slightly higher odds of being partially vs. not vaccinated (OR = 1.04, 95% CI = 1.02–1.07). This same

pattern of effects was found for living with extended family ($p < .001$). Living in multi-person households, relative to living in a single-family household, was associated with greater odds of being fully vaccinated, and lower odds of being partially vs. not vaccinated (p 's < .001), and greater odds of being fully vs. partially vaccinated (OR = 2.0, 95% CI = 1.85–2.17). By contrast, living alone was associated with lower odds of being fully vs. partially (OR = 0.35, 95% CI = 0.33–0.37) and not vaccinated (OR = 0.87, 95% CI = 0.85–0.88).

Area level deprivation was associated with lower odds, and household income higher odds, of being vaccinated (p 's < .001). Those with medium (OR = 1.12, 95% CI = 1.09–1.14) or high mobility (OR = 1.06, 95% CI = 1.00–1.12), had higher odds of being fully vs. partially vaccinated, and lower odds of being partially vs. not vaccinated (OR = 0.91, 95% CI = 0.89–0.93 and OR = 0.85, 95% CI = 0.80–0.90, respectively). Medium and high mobility was associated with slightly greater odds, of being fully vs. not vaccinated (p 's < .001). Many of these effects were small however, and statistical significance may reflect the large sample size but not necessarily meaningful effects.

Discussion

This study demonstrates the presence of inequity even in a country where the overall transmission and impact of COVID-19 was low relative to other nations [13]. Inequities for indigenous (Māori) and ethnic minority (Pacific) groups

in New Zealand were widespread—with Māori and Pacific peoples being more likely to test positive, be hospitalised, and die following positive test results, but generally less likely to be fully vaccinated. Several other variables were predictive of worse outcomes, including disability status (being disabled), high residential mobility, and living in crowded and low-quality housing. Active enrolment with a primary health organisation was a particularly strong predictor of at least partial vaccination. Overall, ethnic group differences in COVID-19 outcomes were smaller when accounting for policy amenable socioeconomic factors.

After adjusting for a range of household structure and socioeconomic variables, Māori remained at 1.2–1.3 times greater odds than non-Māori, and Pacific peoples at 1.3–1.9 times greater odds than non-Pacific peoples of testing positive, being hospitalised, and dying following a positive COVID-19 test. As both Māori and Pacific are at higher risk of worse outcomes, these differences would likely be greater if using a non-Māori/non-Pacific comparison group. This reinforces the need for proactive efforts at the start of a pandemic to prevent the worst outcomes being experienced by indigenous and ethnic minority groups. Official testing data indicate high levels of testing uptake by Māori and Pacific peoples in some age groups [37]. This initial engagement with the COVID-19 response did not translate into correspondingly higher vaccination rates. Although differences were small, vaccination status was assessed at May 2023, nearly two years after the widespread availability of the COVID-19 vaccine in New Zealand. Future research could map out the time course of inequity throughout the pandemic, such as assessing vaccination rates at different points of time (e.g. 2 months, 6 months, 1 year on from their availability).

Those with a disability had 2.3 and 2.2 times the odds of hospitalisation and death than those without a disability, consistent with past research [42]. The effects are notable given disability status was broadly measured in the Census (i.e. having difficulty seeing, hearing, walking or climbing steps, remembering or concentrating, washing all over or dressing, communicating). Stronger effects may have been found for more specific measures of disabilities (e.g. respiratory problems). Yet this also points to more nuanced ways in which people can be more susceptible to worse outcomes. For example, disabled people may have greater difficulty in accessing medical help [43].

In terms of vaccination status, not having an active enrolment with a primary health organisation (i.e. general practice) immediately prior to the vaccination rollout was particularly strongly associated with being unvaccinated. This is notable as the COVID-19 vaccine was freely available in New Zealand during the study period at many non-primary healthcare locations, including pharmacies, local community centres, school and sports facilities. This finding suggests those unenrolled are disconnected from and unlikely to engage with the health system. As such, administrative data on primary healthcare enrolment could be a vital tool for identifying and developing interventions for hard-to-reach populations.

Strengths and limitations

Socioeconomic issues such as high residential mobility, household crowding, and poor-quality housing were associated with worse COVID-19 outcomes. These effects are important as they could partly explain ethnic group differences in COVID-19 outcomes and are policy amenable. Although ethnic group differences remained after accounting for these factors, it is important to note that we focused on adjusting for policy amenable factors that were measured and available in data sources in the IDI. There are likely other relevant factors not measured in this study, such as distance to vaccination sites [22] that, if adjusted for, could alter (either increase or decrease) the associations documented in this study. Future work could build on this initial examination by applying formally hypothesised causal models, including mediation models to study specific causal pathways to COVID-19 outcomes. We also did not test for non-linear effects (for age, income, and deprivation for example). Therefore, our analyses likely miss some nuance in the way these variables are associated with COVID-19 outcomes. For example, the odds of hospitalisation may be relatively similar across younger ages but increase more sharply in much older ages.

This study demonstrates strengths in using linked administrative data to provide crucial pandemic insights, particularly with the availability of high-quality data on COVID-19 health outcomes and relevant demographic and socioeconomic variables. Data quality here was facilitated through mandatory reporting requirements in New Zealand for hospitalisations and deaths, and complete initial reporting of supervised PCR COVID-19 test results. We limited our testing outcome variable to these records, however hospitalisation and death outcomes also made use of self-reported positive RAT results. Thus, there may be some underreporting of positive tests without subsequent hospitalisations or deaths. However, some deaths may also be underreported, particularly given cause of death information was unavailable in the IDI beyond 2018.

The availability of linked 2018 Census data, which was conducted relatively close to the start of the pandemic, was also an important and high-quality source of demographic and socioeconomic data. Nonetheless, limitations with using 2018 Census data remain. It is possible that household characteristics had changed between measurement at the Census in 2018 and the start of the pandemic, which will add error to estimates for household effects. Moreover, it means our models miss COVID-19 outcomes for those not included in the 2018 Census, such as more recent births and migrants, those overseas on census night, and groups undercounted by the census (i.e. Māori and Pacific people) [44]. As this study is retrospective it is informative for ongoing management of COVID-19 and other existing and new health issues. However, greater resourcing could allow such analyses to be conducted near real-time during early and critical periods of a pandemic.

Conclusion

This study demonstrates social inequities present in diverse COVID-19 outcomes in New Zealand—a nation with COVID-19 pandemic outcomes that have been compared very favourably

at the international level. In our whole population analysis we identified several groups, including Māori, Pacific peoples, those with a disability, and those experiencing poor housing conditions (including crowded, low-quality housing, and high residential mobility) who experienced worse health outcomes during the first three years of the COVID-19 pandemic. Results suggest socioeconomic circumstances such as housing conditions may partially explain ethnic group differences in COVID-19 health outcomes. Moreover, the socioeconomic variables assessed here are policy amenable, meaning current and future inequitable health outcomes should not be treated as inevitable or ignored. As the effects of these variables were observed across a broad range of COVID-19 health outcomes, they will be important to consider in current and future pandemic management.

Acknowledgements

Funding for this work was provided New Zealand Ministry of Business, Innovation and Employment COVID-19 Programme Project UOAX1941 and Health Research Council of New Zealand Project Grant 20/1442. Initial funding was also provided by the National Hauora Coalition.

Statement of conflicts of interest

None.

Ethics

Ethics approval was sought but this project was confirmed as out of scope of ethical review by the New Zealand Health and Disability Ethics Committee on 1 July 2021 (21/STH/161).

Disclaimer

These results are not official statistics. They have been created for research purposes from the Integrated Data Infrastructure (IDI) which is carefully managed by Stats NZ. For more information about the IDI please visit <https://www.stats.govt.nz/integrated-data/>. Access to the data used in this study was provided by Stats NZ under conditions designed to give effect to the security and confidentiality provisions of the Data and Statistics Act 2022. The results presented in this study are the work of the authors, not Stats NZ or individual data suppliers.

Data availability

Data used in this study was accessed and analysed within the Integrated Data Infrastructure (IDI). IDI data is not publicly available and cannot be shared by the authors. Access to the data used in this study was provided by Stats NZ under conditions designed to give effect to the security and confidentiality provisions of the Data and Statistics Act 2022. For more information on IDI access see <https://www.stats.govt.nz/integrated-data/apply-to-use-microdata-for-research/>.

References

1. Dessie ZG, Zewotir T. Mortality-related risk factors of COVID-19: a systematic review and meta-analysis of 42 studies and 423,117 patients. *BMC infectious diseases*. 2021 Aug 21;21(1):855. <https://doi.org/10.1186/s12879-021-06536-3>
2. Ho FK, Petermann-Rocha F, Gray SR, Jani BD, Katikireddi SV, Niedzwiedz CL, Foster H, Hastie CE, Mackay DF, Gill JM, O'Donnell C. Is older age associated with COVID-19 mortality in the absence of other risk factors? General population cohort study of 470,034 participants. *PLoS one*. 2020 Nov 5;15(11):e0241824. <https://doi.org/10.1371/journal.pone.0241824>
3. Wolff D, Nee S, Hickey NS, Marschollek M. Risk factors for Covid-19 severity and fatality: a structured literature review. *Infection*. 2021 Feb;49:15-28. <https://doi.org/10.1007/s15010-020-01509-1>
4. United Nations. The impact of COVID-19 on Indigenous peoples. 2020. Policy brief No 70, Department of Economic and Social Affairs. Available at: <http://www.un.org/development/desa/dpad/publication/un-des-a-policy-brief-70-the-impact-of-covid19-on-Indigenous-peoples/>.
5. Mallard A, Pesantes MA, Zavaleta-Cortijo C, Ward J. An urgent call to collect data related to COVID-19 and Indigenous populations globally. *BMJ Global Health*. 2021 Mar 1;6(3):e004655. <https://doi.org/10.1136/bmjgh-2020-004655>
6. Pickering K, Galappaththi EK, Ford JD, Singh C, Zavaleta-Cortijo C, Hyams K, Miranda JJ, Arotoma-Rojas I, Togarepi C, Kaur H, Arvind J. Indigenous peoples and the COVID-19 pandemic: a systematic scoping review. *Environmental Research Letters*. 2023 Feb 13;18(3):033001. <https://doi.org/10.1088/1748-9326/acb804>
7. Agyemang C, Richters A, Jolani S, Hendricks S, Zalpuri S, Yu E, Pijls B, Prins M, Stronks K, Zeegers MP. Ethnic minority status as social determinant for COVID-19 infection, hospitalisation, severity, ICU admission and deaths in the early phase of pandemic: a meta-analysis. *BMJ Glob Health*. 2021; 6 (11): e007433. <https://doi.org/10.1088/1748-9326/acb804>
8. Amele S, Kibuchi E, McCabe R, Pearce A, Henery P, Hailey K, Fagbamigbe AF, Kurdi A, McCowan C, Simpson CR, Dibben C. Ethnic inequalities in positive SARS-CoV-2 tests, infection prognosis, COVID-19 hospitalisations and deaths: analysis of 2 years of a record linked national cohort study in Scotland. *J Epidemiol Community Health*. 2023 Oct 1;77(10):641-8. <https://doi.org/10.1136/jech-2023-220501>
9. Gross CP, Essien UR, Pasha S, Gross JR, Wang SY, Nunez-Smith M. Racial and ethnic disparities in population-level Covid-19 mortality. *Journal of general*

- internal medicine. 2020 Oct;35:3097-9. <https://doi.org/10.1007/s11606-020-06081-w>
10. Joynt Maddox KE, Reidhead M, Grotzinger J, McBride T, Mody A, Nagasako E, Ross W, Steensma JT, Barker AR. Understanding contributors to racial and ethnic inequities in COVID-19 incidence and mortality rates. *Plos one*. 2022 Jan 28;17(1):e0260262. <https://doi.org/10.1371/journal.pone.0260262>
11. Kirby T. Evidence mounts on the disproportionate effect of COVID-19 on ethnic minorities. *The Lancet Respiratory Medicine*. 2020 Jun 1;8(6):547-8. [https://doi.org/10.1016/s2213-2600\(20\)30228-9](https://doi.org/10.1016/s2213-2600(20)30228-9)
12. McLaren J. Racial disparity in COVID-19 deaths: Seeking economic roots with census data. *The BE Journal of Economic Analysis & Policy*. 2021 Jul 14;21(3):897-919. <https://doi.org/10.1515/bejeap-2020-0371>
13. Karlinsky A, Kobak D. Tracking excess mortality across countries during the COVID-19 pandemic with the World Mortality Dataset. *elife*. 2021 Jun 30;10:e69336. <https://doi.org/10.7554/eLife.69336>
14. Gurney J, Stanley J, Sarfati D. The inequity of morbidity: Disparities in the prevalence of morbidity between ethnic groups in New Zealand. *Journal of comorbidity*. 2020 Nov 6;10:2235042X20971168. <https://doi.org/10.1177/2235042X20971168>
15. Te Rangi Hiroa. 'The smallpox epidemic among Maoris in the Northern District'. 1914. *Aust.Med. Con.*, 10th session, Feb., pp.212-24. Cited in Pool DI (1991).
16. Wilson N, Barnard LT, Summers JA, Shanks GD, Baker MG. Differential mortality rates by ethnicity in 3 influenza pandemics over a century, New Zealand. *Emerging infectious diseases*. 2012 Jan;18(1):71. <https://doi.org/10.3201/eid1801.110035>
17. Baker MG, Kvalsvig A, Verrall AJ, Telfar-Barnard L, Wilson N. New Zealand's elimination strategy for the COVID-19 pandemic and what is required to make it work. *The New Zealand Medical Journal (Online)*. 2020 Apr 3;133(1512):10-4.
18. McLeod M, Gurney J, Harris R, Cormack D, King P. COVID-19: we must not forget about Indigenous health and equity. *Australian and New Zealand journal of public health*. 2020 Aug;44(4):253. <https://doi.org/10.1111/1753-6405.13015>
19. Jefferies S, French N, Gilkison C, Graham G, Hope V, Marshall J. COVID-19 in New Zealand and the impact of the national response: a descriptive epidemiological study. *Lancet Public Health*. 2020; 5: e612–23. [https://doi.org/10.1016/S2468-2667\(20\)30225-5](https://doi.org/10.1016/S2468-2667(20)30225-5)
20. Steyn N, Binny RN, Hannah K, Hendy SC, James A, Lustig A, Ridings K, Plank MJ, Sporle A. Māori and Pacific People in New Zealand have higher risk of hospitalisation for COVID-19. *MedRxiv*. 2020 Dec 28;2020-12. <https://doi.org/10.1101/2020.12.25.20248427>
21. Whitehead J, Gan H, Heerikhuisen J, Gray G, Richardson T, Brown P, Lawrenson R. Inequities in COVID-19 Omicron infections and hospitalisations for Māori and Pacific people in Te Manawa Taki Midland region, New Zealand. *Epidemiology & Infection*. 2023 Jan;151:e74. <https://doi.org/10.1017/S0950268823000572>
22. Whitehead J, Carr PA, Scott N, Lawrenson R. Structural disadvantage for priority populations: the spatial inequity of COVID-19 vaccination services in Aotearoa. *NZMA*. 2022 Mar 11;135(1551):1175-8716.
23. Waitangi Tribunal. Haumaru: The Covid-19 priority report. 2021. Wai2575. .
24. Harvey, E., Hobbs, M., Kvalsvig, A., Mackenzie, F., O'Neale, D., & Turnbull, S. A Summary of Risk Factors for COVID-19 Infection in Aotearoa New Zealand. *Covid-19 Modelling Aotearoa preprint*, 2023. <https://www.covid19modelling.ac.nz/a-summary-of-risk-factors-for-covid-19-infection-in-aotearoa-new-zealand/>.
25. Tai DB, Shah A, Doubeni CA, Sia IG, Wieland ML. The disproportionate impact of COVID-19 on racial and ethnic minorities in the United States. *Clinical infectious diseases*. 2021 Feb 15;72(4):703-6. <https://doi.org/10.1093/cid/ciaa815>
26. Labberton AS, Godøy A, Elgersma IH, Strand BH, Telle K, Arnesen T, Nygård KM, Indseth T. SARS-CoV-2 infections and hospitalisations among immigrants in Norway-significance of occupation, household crowding, education, household income and medical risk: a nationwide register study. *Scandinavian Journal of Public Health*. 2022 Aug;50(6):772-81. <https://doi.org/10.1177/14034948221075029>
27. Lassale C, Gaye B, Hamer M, Gale CR, Batty GD. Ethnic disparities in hospitalisation for COVID-19 in England: The role of socioeconomic factors, mental health, and inflammatory and pro-inflammatory factors in a community-based cohort study. *Brain, behavior, and immunity*. 2020 Aug 1;88:44-9. <https://doi.org/10.1016/j.bbi.2020.05.074>
28. Mena GE, Martinez PP, Mahmud AS, Marquet PA, Buckee CO, Santillana M. Socioeconomic status determines COVID-19 incidence and related mortality in Santiago, Chile. *Science*. 2021 May 28;372(6545):eabg5298. <https://doi.org/10.1126/science.abg5298>
29. Ahmad K, Erqou S, Shah N, Nazir U, Morrison AR, Choudhary G, Wu WC. Association of poor housing conditions with COVID-19 incidence and mortality across US counties. *PloS one*. 2020 Nov 2;15(11):e0241327. <https://doi.org/10.1371/journal.pone.0241327>
30. Ghosh AK, Venkatraman S, Soroka O, Reshetnyak E, Rajan M, An A, Chae JK, Gonzalez C, Prince J,

- DiMaggio C, Ibrahim S. Association between overcrowded households, multigenerational households, and COVID-19: a cohort study. *Public Health*. 2021 Sep 1;198:273-9. <https://doi.org/10.1016/j.puhe.2021.07.039>
31. Hu M, Roberts JD, Azevedo GP, Milner D. The role of built and social environmental factors in Covid-19 transmission: A look at America's capital city. *Sustainable Cities and Society*. 2021 Feb 1;65:102580. <https://doi.org/10.1016/j.scs.2020.102580>
32. Mousa A, Winskill P, Watson OJ, Ratmann O, Monod M, Ajelli M, Diallo A, Dodd PJ, Grijalva CG, Kiti MC, Krishnan A. Social contact patterns and implications for infectious disease transmission—a systematic review and meta-analysis of contact surveys. *Elife*. 2021 Nov 25;10:e70294. <https://doi.org/10.7554/eLife.70294>
33. Nafilyan V, Islam N, Ayoubkhani D, Gilles C, Katikireddi SV, Mathur R, Summerfield A, Tingay K, Asaria M, John A, Goldblatt P. Ethnicity, household composition and COVID-19 mortality: a national linked data study. *Journal of the Royal Society of Medicine*. 2021 Apr;114(4):182-211. <https://doi.org/10.1177/0141076821999973>
34. Martin CA, Jenkins DR, Minhas JS, Gray LJ, Tang J, Williams C, Sze S, Pan D, Jones W, Verma R, Knapp S. Socio-demographic heterogeneity in the prevalence of COVID-19 during lockdown is associated with ethnicity and household size: Results from an observational cohort study. *EClinicalMedicine*. 2020 Aug 1;25. <https://doi.org/10.1016/j.eclinm.2020.100466>
35. Gillies CL, Rowlands AV, Razieh C, Nafilyan V, Chudasama Y, Islam N, Zaccardi F, Ayoubkhani D, Lawson C, Davies MJ, Yates T. Association between household size and COVID-19: A UK Biobank observational study. *Journal of the Royal Society of Medicine*. 2022 Apr;115(4):138-44. <https://doi.org/10.1177/01410768211073923>
36. Stats NZ. Integrated Data Infrastructure. 2022. Available from: <https://www.stats.govt.nz/integrated-data/integrated-data-infrastructure/>.
37. Health New Zealand—Te Whatu Ora. COVID-19 case demographics. 2024. Accessed 13 March 2024 <https://www.tewhatauora.govt.nz/our-health-system/data-and-statistics/covid-19-data/covid-19-case-demographics/>.
38. Ministry of Health. (2020). Immunisation Handbook 2020. Wellington: Ministry of Health.
39. Goodyear RK, Fabian A, Hay J. Finding the crowding index that works best for New Zealand. 2011. Working Paper No 11–04. Available from www.stats.govt.nz.
40. 245335844 Atkinson J, Salmond C, Crampton P (2019). NZDep2018 Index of Deprivation, Final Research Report, December 2020. Wellington: University of Otago.
41. Hanley JA, Negassa A, Edwardes MD, Forrester JE. Statistical analysis of correlated data using generalized estimating equations: an orientation. *American journal of epidemiology*. 2003 Feb 15;157(4):364-75. <https://doi.org/10.1093/aje/kwf215>
42. Kavanagh A, Hatton C, Stancliffe RJ, Aitken Z, King T, Hastings R, Totsika V, Llewellyn G, Emerson E. Health and healthcare for people with disabilities in the UK during the COVID-19 pandemic. *Disability and Health Journal*. 2022 Jan 1;15(1):101171. <https://doi.org/10.1016/j.dhjo.2021.101171>
43. Shakespeare T, Ndagire F, Seketi QE. Triple jeopardy: disabled people and the COVID-19 pandemic. *The Lancet*. 2021 Apr 10;397(10282):1331-3. [https://doi.org/10.1016/S0140-6736\(21\)00625-5](https://doi.org/10.1016/S0140-6736(21)00625-5)
44. Stats NZ. Post-enumeration survey: 2018. 2020. Retrieved from www.stats.govt.nz

Abbreviations

IDI:	Integrated Data Infrastructure
MOH:	Ministry of Health



Supplementary Appendices

The Supplementary Appendix contains reduced models that include only individual characteristics variables, and individual and household structure variables, as predictors of COVID-19 infection outcomes (Supplementary Tables 1, 2) and vaccination status (Supplementary Tables 3, 4).

Supplementary Table 1: Individual predictors of COVID-19 infection outcomes

	Positive covid test (vs. negative covid test) for tests to 16 February 2022				Hospitalisation following positive test (vs. not hospitalised) for positive tests to 02 June 2022				Death following positive test (vs. no death) for positive tests to 14 February 2023			
	Odds ratio	95% CI Lower	95% CI Upper	<i>p</i>	Odds ratio	95% CI Lower	95% CI Upper	<i>p</i>	Odds ratio	95% CI Lower	95% CI Upper	<i>p</i>
Māori Ethnicity (0 no, 1 yes)	1.28	1.16	1.41	<.001	1.43	1.34	1.53	<.001	1.54	1.33	1.79	<.001
Pacific Ethnicity (0 no, 1 yes)	2.21	1.95	2.50	<.001	1.98	1.84	2.14	<.001	1.58	1.28	1.95	<.001
Sex (0 women, 1 men)	0.99	0.94	1.06	.844	0.78	0.74	0.81	<.001	1.61	1.49	1.75	<.001
Age	0.90	0.88	0.92	<.001	1.50	1.48	1.53	<.001	3.57	3.43	3.72	<.001
PHO enrolment (0 no, 1 yes)	0.75	0.63	0.90	.002	1.59	1.28	1.97	<.001	0.55	0.36	0.84	.006
Disability status (0 no, 1 yes)	0.85	0.70	1.03	.092	2.90	2.69	3.12	<.001	2.38	2.15	2.63	<.001

Note. $N(\text{test outcome}) = 1,333,944$, $N(\text{hospitalisation outcome}) = 697,614$. $N(\text{death outcome}) = 1,314,810$. * $p < .05$, ** $p < .01$, *** $p < .001$.

Supplementary Table 2: Individual and household structure predictors of COVID-19 infection outcomes

	Positive covid test (vs. negative covid test) for tests to 16 February 2022				Hospitalisation following positive test (vs. not hospitalised) for positive tests to 02 June 2022				Death following positive test (vs. no death) for positive tests to 14 February 2023			
	Odds ratio	95% CI Lower	95% CI Upper	<i>p</i>	Odds ratio	95% CI Lower	95% CI Upper	<i>p</i>	Odds ratio	95% CI Lower	95% CI Upper	<i>p</i>
Māori Ethnicity (0 no, 1 yes)	1.28	1.17	1.39	<.001	1.40	1.31	1.49	<.001	1.49	1.28	1.73	<.001
Pacific Ethnicity (0 no, 1 yes)	2.22	1.99	2.47	<.001	1.93	1.79	2.09	<.001	1.51	1.22	1.88	<.001
Sex (0 women, 1 men)	1.00	0.95	1.05	.959	0.79	0.75	0.83	<.001	1.74	1.59	1.89	<.001
Age	0.90	0.89	0.92	<.001	1.47	1.45	1.50	<.001	3.45	3.31	3.60	<.001
PHO enrolment (0 no, 1 yes)	0.75	0.64	0.89	.001	1.64	1.32	2.04	<.001	0.58	0.38	0.88	.010
Disability status (0 no, 1 yes)	0.87	0.74	1.03	.103	2.80	2.60	3.01	<.001	2.34	2.12	2.59	<.001
Child under 5 (0 no, 1 yes)	1.30	1.21	1.40	<.001	0.99	0.92	1.06	.791	1.03	0.80	1.33	.835
Extended family (0 no, 1 yes)	1.39	1.20	1.60	<.001	1.25	1.13	1.39	<.001	1.30	1.06	1.60	.011
<i>Household composition^a</i>												
Two or more families	1.24	1.08	1.42	.002	1.05	0.94	1.17	.417	1.02	0.81	1.29	.880
Multi-person	1.10	0.97	1.26	.133	1.33	1.18	1.51	<.001	1.43	1.12	1.83	.004
One person	0.83	0.73	0.95	.006	1.54	1.42	1.67	<.001	1.42	1.28	1.58	<.001

Note. $N(\text{test outcome}) = 1,333,944$, $N(\text{hospitalisation outcome}) = 697,614$. $N(\text{death outcome}) = 1,314,810$. * $p < .05$, ** $p < .01$, *** $p < .001$.

^aReference category is those in single family households.

Supplementary Table 3: Individual predictors of vaccination status as at 10 May 2023

	Fully vaccinated (vs. Partially vaccinated)				Fully vaccinated (vs. Not vaccinated)				Partially vaccinated (vs. Not vaccinated)			
	Odds ratio	95% CI Lower	95% CI Upper	<i>p</i>	Odds ratio	95% CI Lower	95% CI Upper	<i>p</i>	Odds ratio	95% CI Lower	95% CI Upper	<i>p</i>
Māori Ethnicity (0 no, 1 yes)	0.67	0.66	0.69	<.001	0.70	0.70	0.71	<.001	0.85	0.83	0.88	<.001
Pacific Ethnicity (0 no, 1 yes)	0.83	0.81	0.86	<.001	0.84	0.82	0.85	<.001	1.02	0.99	1.06	.214
Sex (0 women, 1 men)	0.92	0.91	0.94	<.001	1.20	1.19	1.21	<.001	1.06	1.04	1.08	<.001
Age	2.75	2.71	2.79	<.001	1.13	1.13	1.13	<.001	0.64	0.64	0.64	<.001
PHO enrolment (0 no, 1 yes)	0.97	0.93	1.00	.070	11.18	11.05	11.30	<.001	6.68	6.44	6.94	.057
Disability status (0 no, 1 yes)	0.71	0.68	0.74	<.001	0.69	0.68	0.70	<.001	1.05	1.00	1.10	<.001

Note. $N(\text{Fully vs. partially}) = 2,810,535$. $N(\text{Fully vs. none}) = 3,060,945$. $N(\text{Partially vs. none}) = 359,766$. * $p < .05$, ** $p < .01$, *** $p < .001$.

Supplementary Table 4: Individual and household structure predictors of vaccination status as at 10 May 2023

	Fully vaccinated (vs. Partially vaccinated)				Fully vaccinated (vs. Not vaccinated)				Partially vaccinated (vs. Not vaccinated)			
	Odds ratio	95% CI Lower	95% CI Upper	<i>p</i>	Odds ratio	95% CI Lower	95% CI Upper	<i>p</i>	Odds ratio	95% CI Lower	95% CI Upper	<i>p</i>
Māori Ethnicity	0.71	0.69	0.72	<.001	0.72	0.71	0.72	<.001	0.85	0.83	0.87	<.001
Pacific Ethnicity	0.92	0.89	0.95	<.001	0.87	0.85	0.89	<.001	1.01	0.97	1.04	.075
Sex (0 women, 1 men)	0.91	0.89	0.93	<.001	1.20	1.19	1.21	<.001	1.06	1.04	1.08	<.001
Age	2.79	2.75	2.83	<.001	1.13	1.12	1.13	<.001	0.64	0.64	0.65	<.001
PHO enrolment (0 no, 1 yes)	1.01	0.97	1.05	.798	11.60	11.47	11.74	<.001	6.60	6.36	6.86	<.001
Disability status (0 no, 1 yes)	0.70	0.67	0.74	<.001	0.69	0.68	0.70	<.001	1.05	1.00	1.10	.038
Child under 5 (0 no, 1 yes)	0.41	0.40	0.42	<.001	0.63	0.62	0.63	<.001	1.03	1.00	1.05	.026
Extended family (0 no, 1 yes)	0.91	0.87	0.95	<.001	0.81	0.79	0.83	<.001	1.12	1.07	1.17	<.001
<i>Household composition^a</i>												
Two or more families	1.11	1.06	1.16	<.001	1.17	1.14	1.20	<.001	0.96	0.92	1.01	<.093
Multi-person	1.87	1.73	2.02	<.001	1.32	1.29	1.36	<.001	0.60	0.56	0.65	<.001
One person	0.32	0.31	0.34	<.001	0.77	0.76	0.79	<.001	0.99	0.93	1.05	.748

Note. $N(\text{Fully vs. partially}) = 2,810,535$. $N(\text{Fully vs. none}) = 3,060,945$. $N(\text{Partially vs. none}) = 359,766$. * $p < .05$, ** $p < .01$, *** $p < .001$.

^aReference category is those in single family households.

