Sensitivity of RT-PCR tests for SARS-CoV-2 through time

Rachelle N Binny[1,2], Patricia Priest[3], Nigel P French[4], Matthew Parry[2,5], Audrey Lustig[1,2], Shaun C Hendy[2,6], Oliver J Maclaren[7], Kannan M Ridings[2,6], Nicholas Steyn[2,6,8], Giorgia Vattiato[2,6,9], Michael J Plank[2,9,*]

[1]Manaaki Whenua-Landcare Research, Lincoln, New Zealand.
[2]Te Pūnaha Matatini, Centre of Research Excellence in Complex Systems, New Zealand.
[3]Department of Preventive and Social Medicine, Dunedin School of Medicine, University of Otago, Dunedin, New Zealand.
[4]Tāwharau Ora/School of Veterinary Science, Massey University, Palmerson North, New Zealand.
[5]Department of Mathematics and Statistics, University of Otago, Dunedin, New Zealand.
[6]Department of Physics, University of Auckland, Auckland, New Zealand.
[7]Department of Engineering Science, University of Auckland, Auckland, New Zealand.
[8]Department of Statistics, University of Oxford, Oxford, UK.
[9]School of Mathematics and Statistics, University of Canterbury, New Zealand.

[*]Corresponding author. Email: michael.plank@canterbury.ac.nz

Running title: RT-PCR test sensitivity for SARS-CoV-2

ORCID IDs:

Rachelle N. Binny: 0000-0002-3433-0417

Patricia Priest: 0000-0003-2311-6236

Nigel French: 0000-0002-6334-0657

Matthew Parry: 0000-0002-6588-0219

Audrey Lustig: 0000-0002-0869-3847

Shaun C. Hendy: 0000-0003-3468-6517

Oliver Maclaren: 0000-0002-3982-4539

Kannan M Ridings: 0000-0002-8801-4428

Nicholas Steyn: 0000-0001-8904-2941

Giorgia Vattiato: 0000-0002-7994-7541

Michael J. Plank: 0000-0002-7539-3465

Abstract

Reverse transcriptase polymerase chain reaction (RT-PCR) tests are the gold standard for detecting recent infection with SARS-CoV-2. RT-PCR sensitivity varies over the course of an individual's infection, related to changes in viral load. Differences in testing methods, and individual-level variables such as age, may also affect sensitivity. Using data from New Zealand, we estimate the time-varying sensitivity of SARS-CoV-2 RT-PCR under varying temporal, biological and demographic factors. Sensitivity peaks 4-5 days post-infection at 92.7% [91.4%, 94.0%] and remains over 88% between 5 and 14 days post-infection. After the peak, sensitivity declined more rapidly in vaccinated cases compared to unvaccinated, females compared to males, those aged under 40 compared to over 40s, and Pacific peoples compared to other ethnicities. RT-PCR remains a sensitive technique and has been an effective tool in New Zealand's border and post-border measures to control COVID-19. Our results inform model parameters and decisions concerning routine testing frequency.

Keywords: Test sensitivity; COVID-19; SARS-CoV-2; Reverse transcription-polymerase chain reaction; Surveillance; Pre-symptomatic infections; False negative

Introduction

Reverse transcriptase polymerase chain reaction (RT-PCR) testing is the gold standard worldwide for detecting whether a person has been recently infected with SARS-CoV-2 and nasopharyngeal RT-PCR was the main type of test employed in Aotearoa New Zealand prior to the arrival of the B.1.1.529 (Omicron) variant in 2022. The test can detect the presence of viral RNA in samples, though its sensitivity (the proportion of tests on infected individuals that return a positive result) varies with time as the amount of virus particles shed by an individual (the viral load) changes over the course of their infection [1, 2]. RT-PCR positivity does not necessarily mean an individual is infectious, particularly more than 10-14 days after infection, as non-viable viral RNA may be present for some time [3, 4]. RT-PCR has high specificity for SARS-CoV-2, estimated at close to 100% [5].

Information about how the sensitivity of RT-PCR tests for detecting SARS-CoV-2 varies with time since infection is important to inform case management, optimise test timing with respect to time of exposure or symptom onset, and to parameterise models of the effectiveness of surveillance testing under different testing regimes (see e.g. [6, 7]). However, available data on RT-PCR sensitivity is limited, particularly for the incubation period prior to symptom onset, as testing is frequently triggered by onset of symptoms, and because there is no independent gold standard assay. Datasets are often subject to sampling bias or other biases, therefore analysis of datasets collected under different testing regimes, on different populations, and at different stages of the pandemic is valuable.

Kucirka et al [8] reported the proportion of false-negative results (i.e. proportion of cases that are not detected, equivalent to $1 -$ sensitivity) for RT-PCR tests up to 3 weeks after infection, with a minimum false-negative rate of 20% [12%, 30%] at 8 days after infection (assuming a fixed 5-day incubation period). However, there was high uncertainty in estimates, especially

prior to symptom onset, owing to the relatively small size (1330 tests) of their dataset, pooled from 7 published studies with heterogeneous designs and different sample collection methods. Using a similar modelling approach, Zhang et al [9] reported false-negative rates for up to one week after symptom onset (test data were limited at later times) for 60 symptomatic individuals in Shenzhen, China. The false-negative rate was 100% at 5 days before symptom onset, again with large uncertainty around pre-onset estimates, and decreased to a minimum of 43% on day 7 after the last exposure to an index case. Another study by Hellewell et al [10] used a piecewise regression model to analyse 241 test results from routine testing on 27 UK healthcare workers and estimated a peak RT-PCR sensitivity of 77% [54%, 88%] on day 4 after infection. Date of symptom onset was not recorded for individuals so, for each worker, a time of infection was inferred by assuming symptom onset occurred within a known censored interval, between the time of the last negative asymptomatic test before symptoms developed and the first symptomatic positive test.

Between June 2020 and November 2021, Aotearoa New Zealand routinely tested all international arrivals on day 3 and day 12 after arrival, during a 14-day mandatory stay in government-managed isolation and quarantine (MIQ) facilities. A negative RT-PCR result on day 12 and medical examination was required to exit MIQ. From January 2021, most arrivals were also tested on day 0. Frontline border workers were also routinely tested for much of this period. In addition, extensive contact tracing and testing were conducted during community outbreaks in March-May 2020 (original virus strain [11]), August-September 2020, February-March 2021 and in August-September 2021 (Delta variant [12]). Together with recording of the date of symptom onset and information on age, sex, comorbidities and vaccination status, this provides a rich dataset for inferring information about the characteristics of RT-PCR tests.

In this work, we analyse New Zealand's testing data to estimate the time-varying sensitivity of RT-PCR tests for detecting SARS-CoV-2. We adapt the Bayesian model of Hellewell et al [10] to infer the parameters of a function representing the probability of a positive test result at a given time t after infection. This function is suitable for use as a modelling input. We assess whether sensitivity is affected by a range of temporal, biological and demographic factors.

Methods

Data

Testing data was obtained from the New Zealand Ministry of Health, containing records for 12,615 SARS-CoV-2 RT-PCR tests performed between 26 February 2020 and 30 September 2021 on 4273 unique individuals who were eventually classified as 'confirmed' or 'probable' cases. Samples were collected by nasopharyngeal swab administered by healthcare professionals. Test results reported were "detected" ($n = 4883$ tests), "not detected" ($n = 7618$), "NA" ($n = 71$), "further analysis required" ($n = 19$), "referred" ($n = 4$), "inadequate" ($n = 1$), "indeterminate" ($n = 2$), "to follow" ($n = 9$), or "see comment" (n = 8). Results that were not either "detected" or "not detected" were discarded, leaving 12,501 test results on 4196 unique individuals.

The testing dataset was merged with data from the national COVID-19 surveillance database, EpiSurv (maintained by the Institute of Environmental Science and Research), using a unique patient identifier. Of the 4196 cases, 194 that were classed as "historical" and 676 that never developed symptoms were excluded. Where the symptom onset date differed between datasets, we prioritised the testing dataset. We excluded tests that were conducted either more than 21 days prior to the symptom onset date (n = 1387) or 35 days after symptom onset (n =

2633). Finally, after reviewing preliminary results for tests conducted prior to 15 June 2020 (see Supplementary Data, 'Analysis of testing by time period') we excluded a further 2126 tests on 1384 cases from this period. This left a dataset consisting of 3599 test results for 1888 unique cases (see Table 1 and Figure 1). Out of these cases, 249 were only tested prior to developing symptoms and 18 never received a positive test result (one of these was a 'probable' case, and the remainder were 'confirmed' cases with an excluded positive test result more than 21 days before or 35 days after symptom onset). See Supplementary Data for details of data sources.

Statistical analysis

We adapted the logistic piecewise regression model of Hellewell et al [10] to estimate the probability of testing positive by RT-PCR test (i.e. sensitivity) as a function of time since infection, given a known time of symptom onset. We assume RT-PCR has 100% specificity. The model jointly infers a time of infection T_i^I for each individual case *i* based on their known time of symptom onset T_i^S and unknown incubation period, $T_i^S - T_i^I$. For each individual's incubation period, we used a log-normal prior distribution with mean 5.5 days and s.d. 2.4 days [13]. We accounted for uncertainty in the distribution of incubation periods by conducting a sensitivity analysis where we re-fitted the model using parameters at the upper and lower ends of the confidence intervals estimated by Lauer et al [13] (Supplementary Figure S1). For individuals who tested positive prior to symptom onset, this distribution was truncated such that their time of infection T_i^I must have occurred prior to their first positive test result (i.e. by placing a lower bound on the distribution equal to the number of days between earliest positive result and onset).

For a result $Y_{i,n}$ ($Y_{i,n}$=1 for positive; $Y_{i,n}$=0 for negative) of a test conducted on individual i at time t_n, we model the probability of testing positive $\theta_{i,n}$ as:

$$Y_{i,n} \sim \text{Bernoulli}(\theta_{i,n}),$$

$$\text{logit}(\theta_{i,n}) = \beta_0 + \beta_1(x - C)^2 + (-\beta_1 + \beta_2)(x - C)^2 H(x - C),$$

$$x = t_n - T_i^I,$$

where x is the time between infection and testing, the breakpoint C is the time when the peak in sensitivity occurs, and $H(s)$ is the Heaviside step function that equals 0 if $s < 0$ (i.e. for times x to the left of the breakpoint) and equals 1 if $s > 0$ (i.e. right of the breakpoint). This parameterisation is similar to the piecewise logistic regression of Hellewell et al [10] but the additional quadratic term allows for a smooth peak which provides a better fit to the data. We used moderately informative priors for coefficients $\beta_0 \sim N(0,25)$ and $\beta_1, \beta_2 \sim N(0,1)$, with the latter truncated with an upper bound at 0 so that prior samples of β_1, β_2 are negative. For the breakpoint we used a prior $C \sim N(5,25)$, truncated so that C has a lower bound of 0.

We adapted the model code published by Hellewell et al [10] and fitted to the testing dataset in R 4.1.0 [14] using the rstan package [15]. Samples were drawn from the model using 4 Monte Carlo Markov chains, with a warmup of 1000 iterations followed by 7000 iterations post-warmup for each chain. We assessed convergence using the R hat diagnostic and by visual assessment of trace plots. Data and code to reproduce the results are available at https://github.com/michaelplanknz/pcr-sensitivity-sars-cov-2.

We fit the model to the full test dataset between 15 June 2020 and 30 September 2021 to assess RT-PCR test sensitivity over time since infection. The data were then stratified to compare sensitivity for different subgroups. Cases were grouped based on their vaccination status ('vaccinated' or 'unvaccinated' at time of diagnosis), where they acquired the infection

('overseas' or 'domestic'), age category ('aged 40 yrs or less', or 'over 40 yrs'), gender ('female' or 'male'), whether they had reported comorbidities ('at least one' or 'none'), and ethnicity ('Māori', 'Pacific peoples' or 'other'). The model was re-fit to each data subset to compare RT-PCR test sensitivity between groups. Note, this approach does not account for confounding or interactions between variables or unobserved covariates, so if there is uneven representation of other factors that affect test sensitivity this may bias group estimates. COVID-19 vaccinations (Pfizer-BioNTech) became available in New Zealand in February 2021, starting with frontline workers and at-risk individuals such as those living in aged residential care, and 91% of vaccinated cases were tested after 1 July 2021 when the Delta variant was prevalent (Supplementary Table S2). To reduce confounding, we therefore only consider vaccination status for cases with symptom onset between 1 July 2021 and 30 September 2021, and exclude cases who had vaccination status 'Not Applicable' (m=14) or 'Unknown' (m=65) from this part of the analysis.

Results

Summary data for the 1888 cases with symptom onset between June 2020 and September 2021 is shown in Table 1. An initial empirical estimate of RT-PCR test sensitivity over days since symptom onset (Figure 2A), calculated as the proportion of all tests that were positive over time, showed considerable variation in sensitivity over the course of infection. Sensitivity was 0% (95% confidence interval 0%-6%) around 8 days before symptom onset and increased to 86% (83%-89%) by the day of onset. Sensitivity peaked at 95% (90%-98%) around 5 days after onset, and remained over 85% for the approximately 10 days after onset, gradually declining thereafter. Fitting the logistic regression model to the full test dataset resulted in a posterior probability of testing positive θ (i.e. RT-PCR sensitivity) over time

since infection that was a good visual match to the empirical distribution. Median posterior RT-PCR sensitivity increased from 0% (95% credible interval CI, 0%-0%) on the day of infection, to 48% (30%-64%) at 3 days after infection, and reached a peak of 93% (91%-94%) at 4 days after infection (Figure 2B; Supplementary Table S1). Sensitivity remained high, at more than 88%, for up to 14 days from infection before declining. A sensitivity analysis using different parameters for the incubation period distribution (median 4.7 days or 5.4 days, compared to 5.1 days in the primary analysis) had only a minimal impact on our results (Supplementary Figure S2).

Figure 3 shows that median test sensitivity was high (>87%) in all groups, for the period from around 5 to 14 days from infection; the most likely period of infectiousness. Peak median sensitivity of at least 91% was reached between days 4 and 7, and did not vary importantly between subgroups of vaccination status, source, gender, age, the presence of comorbidities, or ethnicity. However the rate of decline of sensitivity – which may be related to the decline in viral load and shedding – did vary. Sensitivity declined faster in vaccinated individuals, community cases relative to overseas-acquired cases, females, and younger people. It also declined slightly faster in Pacific peoples compared to non-Māori/non-Pacific ethnicities. Supplementary Table S2 and Figures S4-S5 shows detail of the characteristics of the different groups. Summary statistics for all temporal profiles of sensitivity are provided in Supplementary Data.

Discussion

New Zealand's SARS-CoV-2 surveillance testing of border arrivals, workers in MIQ and cases during community outbreaks, offers a valuable opportunity to assess how the performance of RT-PCR tests varies with time since infection and the effects of different risk

factors. Compared to previous studies, the large size of our dataset, and the existence of multiple sequential test results for the same individual, provides greater certainty in estimates of time-varying RT-PCR sensitivity over a longer period of time since infection, including the period prior to symptom onset. During large outbreaks, testing capacity (e.g. limits on lab processing of RT-PCR assays) thresholds may be exceeded and optimising the timings of tests allows more efficient use of finite resources. We estimate that RT-PCR sensitivity peaks at 93% (95% CI, 91%-94%) 4 days after infection (i.e. 1 day prior to symptom onset, assuming an average 5-day incubation period). At symptom onset, median RT-PCR sensitivity is still at 93% and remains over 88% for up to 14 days after infection (or up to 9 days from the average symptom onset, after which time individuals are unlikely to remain infectious).

The estimated timing of peak sensitivity of 4 days after infection falls within the range estimated in other studies, being most similar to Hellewell et al [10], and aligns with the likely timing of peak viral load in the respiratory tract [1, 2]. However, our results suggest that RT-PCR sensitivity peaks higher and is maintained over a longer duration compared to previous estimates [8-10] (Supplementary Figure S3). One possible explanation for this is that New Zealand's elimination strategy and very low prevalence meant that higher cycle threshold (Ct) values were used to define positive results. Samples were typically tested for 35-40 cycles [16] and there are many cases in the data with a Ct value >35 noted. However, data on the cycle threshold (Ct) value were not available in a consistent format (recorded inconsistently as freeform text and only linked to cases, not tests) so we were unable to investigate the quantitative relationship between time since infection and Ct value. We found RT-PCR sensitivity can remain non-negligible for up to 45 days after infection, which is within the maximum shedding duration of 83 days reported by Cevik et al [17] for the upper respiratory tract. While viral RNA may persist at high enough levels to be detectable by RT-

PCR at these later times, it is unlikely to be RNA from live virus [17] meaning individuals are no longer infectious and the test is instead detecting recent infection.

Identifying cases as early as possible, ideally prior to symptom onset, is critical for trace-test-isolate measures to be effective. In line with previous studies, RT-PCR was relatively insensitive (<50%) at detecting SARS-CoV-2 from 0 to 3 days after infection but rapidly increased to relatively high sensitivity (>90%) by 4 days after infection, prior to the average time of symptom onset at 5 days. If contact tracing can identify close contacts of cases while they are still in their incubation period, this suggests there is a reasonable chance of early detection by RT-PCR, allowing rapid isolation of confirmed cases to reduce the risk of onward transmission. In addition, our results show that an RT-PCR-negative sample collected in the first 0-3 days after exposure to an infectious person is not a strong indicator of the absence of infection and further follow-up testing may be required.

Peak sensitivity varied little between the different groups that we analysed, however we found some interesting differences in the temporal profile of RT-PCR test sensitivity. For infections after 1 July 2021, when Delta was prevalent, vaccinated cases had slightly lower median sensitivity than unvaccinated in the early days of infection but we found no meaningful difference in peak sensitivity. After the peak, sensitivity remained high over the period in which all individuals would be expected to clear their infection, though there was some evidence that after this period sensitivity declined slightly faster for vaccinated individuals. These results are consistent with previous findings of similar peak viral loads but a faster rate of viral load decline (i.e. faster viral clearance time) in vaccinated compared to unvaccinated cases [18-20].

The data on overseas cases represents a well-defined cohort who were routinely tested on days 3 and 12 (and day 0/1 from January 2021). This is ideal for estimating sensitivity over

time since infection as there are multiple test results per person and likely a very low percentage of infections were missed. In contrast, New Zealand's community cases were slightly less likely to have multiple tests. This could potentially bias estimates of sensitivity in community cases upwards because any infected individuals who had a single test and returned a false-negative result are, by definition, not represented in the dataset. However, our model estimated similar sensitivity profiles for overseas and community cases, potentially reflecting highly effective contact tracing and high community case ascertainment rates.

Sensitivity declined at a slightly faster rate in females than males, and in those aged under 40 compared to over 40s. This could be correlated with differences in viral load, though findings on associations with gender and age from previous studies are inconsistent. Similar to our results, other studies have found faster rates of viral load decline in younger age groups [21]. Large studies with frequent sequential sampling of viral load have detected a slight increase in peak viral load with age, though differences were not always clinically significant [20, 22, 23] so it is possible our age groupings were too broad to detect age-dependent differences. In contrast, Mahallawi et al [24] found higher viral loads in females compared to males but no association with age, while others have reported no clear differences for gender or age [25]. Sensitivity did not appear to be associated with presence of comorbidities. Māori and Pacific peoples have higher infection fatality rates [26] and higher risk of hospitalisation [27] from COVID-19 compared to non-Māori/non-Pacific people. We found little difference in the sensitivity of RT-PCR tests for detecting infection for these three ethnicity groups. Though sensitivity declined slightly faster for Pacific peoples compared to non-Māori/non-Pacific, it remained high over the critical period in which individuals are likely to be contagious.

Our results may generalize to SARS-CoV-2 infections in other populations or at other times if testing methods (nasopharyngeal swab by trained health professional, and criteria for declaring a positive result) and viral load dynamics of individuals are generally consistent

with this study. RT-PCR sensitivity may differ for populations with different viral dynamics (for example, due to different demographics, infection by other variants, or extent of infection- or vaccine-acquired immunity), however the qualitative trends observed in the group comparisons are still likely to apply.

Our study has some limitations. There may be considerable individual heterogeneity in RT-PCR sensitivity that our model does not consider, for example due to individual variation in viral shedding. Our assumed prior distribution for incubation period with median 5.1 days [13] was based on a study of the original SARS-CoV-2 strain and may not be representative of incubation periods for other variants. The incubation period of the Delta variant, for example, has been estimated at a shorter median of 4.0 days (SD 1.9) [28]. However, our results were relatively insensitive to using a shorter median incubation period of 4.7 days (Supplementary Figure S2). An unavoidable limitation of our analysis is that the dataset by definition excludes any infected individuals who returned false negatives from all tests (except for one probable case). This is partly mitigated by repeat testing on any individuals reducing the likelihood of multiple false negatives, but could potentially bias our estimates of sensitivity upwards. We were unable to include asymptomatic cases in our cohort as a time of symptom onset was required to infer a likely time of infection for each case. Nonetheless, our results may still inform the optimal timing for testing asymptomatic individuals after a possible exposure event if peak viral loads are similar to symptomatic infections, as has been previously suggested [29].

In conclusion, we find that RT-PCR testing remains a sensitive technique for detecting SARS-CoV-2 infection and has proven to be an effective tool in New Zealand's border measures and test-trace-isolate-quarantine approach to COVID-19 prevention and control. However, RT-PCR has its limitations and a negative test result does not rule out the possibility of infection with SARS-CoV-2, particularly for tests conducted in the early days

following infection and prior to onset of symptoms. If clinical suspicion remains high, or if accuracy is important for case management or disease control, then it may be advisable to keep precautionary measures in place and conduct further testing.

Acknowledgements

The authors acknowledge the support of StatsNZ, the Institute of Environmental Science and Research, and the New Zealand Ministry of Health in supplying data in support of this work. The analyses made use of code published by Hellewell et al (2021). We acknowledge the use of New Zealand eScience Infrastructure (NeSI) high performance computing facilities and consulting support as part of this research. New Zealand's national facilities are provided by NeSI and funded jointly by NeSI's collaborator institutions and through the Ministry of Business, Innovation & Employment's Research Infrastructure programme (https://www.nesi.org.nz).

Conflict of interest statement

This research was funded by the New Zealand Government.

Funding statement

This research was funded by the New Zealand Government.

Corresponding author contact information

Michael J Plank. *Email:* michael.plank@canterbury.ac.nz. *Address:* School of Mathematics and Statistics, University of Cantebury, Private Bag 4800, Christchurch 8140, New Zealand.

References

1. He X, Lau EHY, Wu P, et al. Temporal dynamics in viral shedding and transmissibility of COVID-19. Nat Med **2020**; 26:672-5.

2. Singanayagam A, Patel M, Charlett A, et al. Duration of infectiousness and correlation with RT-PCR cycle threshold values in cases of COVID-19, England, January to May 2020. Euro Surveill **2020**; 25:2001483.

3. Owusu D, Pomeroy MA, Lewis NM, et al. Persistent SARS-CoV-2 RNA shedding without evidence of infectiousness: a cohort study of individuals with COVID-19. J Infect Dis **2021**; 224:1362-71.

4. van Kampen JJA, van de Vijver DAMC, Fraaij PLA, et al. Duration and key determinants of infectious virus shedding in hospitalized patients with coronavirus disease-2019 (COVID-19). Nat Comm **2021**; 12:267.

5. Skittrall JP, Wilson M, Smielewska AA, et al. Specificity and positive predictive value of SARS-CoV-2 nucleic acid amplification testing in a low-prevalence setting. Clin Microbiol Infect **2021**; 27:469.e9-.e15.

6. [Preprint] Quilty BJ, Russell TW, Clifford S, et al. Quarantine and testing strategies to reduce transmission risk from imported SARS-CoV-2 infections: a global modelling study. medRxiv **2021**:2021.06.11.21258735.

7. Steyn N, Plank MJ, James A, Binny RN, Hendy SC, Lustig A. Managing the risk of a COVID-19 outbreak from border arrivals. J R Soc Interface **2021**; 18:20210063.

8. Kucirka LM, Lauer SA, Laeyendecker O, Boon D, Lessler J. Variation in false-negative rate of reverse transcriptase polymerase chain reaction–based SARS-CoV-2 tests by time since exposure. Ann Intern Med **2020**; 173:262-7.

9. Zhang Z, Bi Q, Fang S, et al. Insight into the practical performance of RT-PCR testing for SARS-CoV-2 using serological data: a cohort study. Lancet Microbe **2021**; 2:e79-e87.

10. Hellewell J, Russell TW, Matthews R, et al. Estimating the effectiveness of routine asymptomatic PCR testing at different frequencies for the detection of SARS-CoV-2 infections. BMC Med **2021**; 19:106.

11. Jefferies S, French N, Gilkison C, et al. COVID-19 in New Zealand and the impact of the national response: a descriptive epidemiological study. Lancet Public Health **2020**; 5:e612-e23.

12. Jelley L, Douglas J, Ren X, et al. Genomic epidemiology of Delta SARS-CoV-2 during transition from elimination to suppression in Aotearoa New Zealand. Nat Comm **2022**; 13:4035.

13. Lauer SA, Grantz KH, Bi Q, et al. The incubation period of Coronavirus Disease 2019 (COVID-19) from publicly reported confirmed cases: estimation and application. Ann Intern Med **2020**; 172:577-82.

14. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria, **2021**.

15. Stan Development Team. RStan: The r Interface to Stan, **2020**.

16. Ministry of Health. Responses to Official Information Act requests. Information on CT value in Covid-19 RT-PCR testing. Available at: https://www.health.govt.nz/system/files/documents/information-release/h202109044_response.pdf. Accessed 19 May 2022.

17. Cevik M, Tate M, Lloyd O, Maraolo AE, Schafers J, Ho A. SARS-CoV-2, SARS-CoV, and MERS-CoV viral load dynamics, duration of viral shedding, and infectiousness: a systematic review and meta-analysis. Lancet Microbe **2021**; 2:e13-e22.

18. Chia PY, Ong SWX, Chiew CJ, et al. Virological and serological kinetics of SARS-CoV-2 Delta variant vaccine breakthrough infections: a multicentre cohort study. Clin Microbiol Infect **2022**; 28:612.e1-.e7.

19. Kissler SM, Fauver JR, Mack C, et al. Viral dynamics of SARS-CoV-2 variants in vaccinated and unvaccinated persons. N Engl J Med **2021**; 385:2489-91.

20. Singanayagam A, Hakki S, Dunning J, et al. Community transmission and viral load kinetics of the SARS-CoV-2 delta (B.1.617.2) variant in vaccinated and unvaccinated individuals in the UK: a prospective, longitudinal, cohort study. Lancet Infect Dis **2022**; 22:183-95.

21. Néant N, Lingas G, Le Hingrat Q, et al. Modeling SARS-CoV-2 viral kinetics and association with mortality in hospitalized patients from the French COVID cohort. Proc Natl Acad Sci U S A **2021**; 118:e2017962118.

22. Euser S, Aronson S, Manders I, et al. SARS-CoV-2 viral-load distribution reveals that viral loads increase with age: a retrospective cross-sectional cohort study. Int J Epidemiol **2021**; 50:1795-803.

23. Jones TC, Biele G, Mühlemann B, et al. Estimating infectiousness throughout SARS-CoV-2 infection course. Sci **2021**; 373:eabi5273.

24. Mahallawi WH, Alsamiri AD, Dabbour AF, Alsaeedi H, Al-Zalabani AH. Association of viral load in SARS-CoV-2 patients with age and gender. Front Med **2021**; 8:608215.

25. Boan P, Jardine A, Pryce TM. Clinical associations of SARS-CoV-2 viral load using the first WHO International Standard for SARS-CoV-2 RNA. Pathol **2022**; 54:344-50.

26. Steyn N, Binny RN, Hannah K, et al. Estimated inequities in COVID-19 infection fatality rates by ethnicity for Aotearoa New Zealand. N Z Med J **2020**; 133:28-39.

27. Steyn N, Binny RN, Hannah K, et al. Māori and Pacific people in New Zealand have a higher risk of hospitalisation for COVID-19. N Z Med J **2021**; 134:28-43.

28. Zhang M, Xiao J, Deng A, et al. Transmission dynamics of an outbreak of the COVID-19 Delta variant B.1.617.2 - Guangdong Province, China, May-June 2021. China CDC Wkly **2021**; 3:584-6.

29. Zuin M, Gentili V, Cervellati C, Rizzo R, Zuliani G. Viral load difference between symptomatic and asymptomatic COVID-19 patients: systematic review and meta-analysis. Infect Dis Rep **2021**; 13:645-53.

Figure captions

Figure 1. Example data for 50 of the cases in the dataset. Blue circles represent negative tests, red crosses represent positive tests plotted against time relative to symptom onset on the horizontal axis. To aid visual interpretation, blue and red lines represent periods of time between tests when the most recent test result was respectively negative or positive.

Figure 2. RT-PCR test sensitivity to detect SARS-CoV-2 infection over time since symptom onset (A) and infection (B). A. Raw data on the proportion of tests that are positive (black dotted) against time relative to symptom onset for 3599 test results for 1888 unique cases. Grey shaded region shows the 95% binomial confidence interval**. B.** Posterior median (blue solid) and 95% credible interval (blue shaded region) for the probability of testing positive θ from the logistic regression; and empirical mean (black dotted) and 95% uncertainty interval (grey shaded region) of the empirical distribution, calculated from the posterior sample of the times of infection T_i^I.

Figure 3. RT-PCR test sensitivity over days since time of infection for different groups by vaccination status, source of infection, gender, age group, presence of comorbidities, and ethnicity. Median (solid lines) and 95% credible interval (shaded regions) for the probability of testing positive θ and empirical mean (dotted line).

Table 1. Characteristics of $m = 1888$ cases with symptom onset between 15 June 2020 and 30 September 2021, who were tested by RT-PCR at least once between 3 weeks prior to symptom onset and 5 weeks after onset

Variable	No. of cases (*m*)	% of total cases
Sex		
Female	932	49.4%
Male	955	50.6%
Unknown	1	0.1%
Age (years)		
0-19	550	29.1%
20-39	765	40.5%
40-59	459	24.3%
60-79	109	5.8%
>=80	5	0.3%
Mean 31 (IQR 17-44)	*1888*	*100%*
Status		
Confirmed	1887	99.9%
Probable	1	0.1%
Overseas-acquired		
Yes	610	32.3%
No	1271	67.3%
Unknown or NULL	7	0.4%
Time period (variant)		

Symptom onset between 15 June 2020 – 30 June 2021 (original strain or earlier variants of concern)	670	35.5%
Symptom onset after 1 July 2021 (Delta variant)	1218	64.5%
Comorbidities		
At least one	266	14.1%
None	1622	85.9%
Vaccinated at time of diagnosis (at least one dose)		
Yes	242	12.8%
No[a]	1510	80.0%
Unknown	136	7.2%
Ethnicity		
Māori	225	11.9%
Pacific Peoples	844	44.7%
Non-Māori and non-Pacific ('Other')	811	43.0%
Unknown	8	0.4%
Number of tests		
1	1730	52.9%
2	981	30.0%
3	353	10.8%
4	127	3.9%
5	49	1.5%
>=6	32	1.0%

Days from symptom onset to first test		
<-5	347	10.6%
-5	69	2.1%
-4	111	3.4%
-3	127	3.9%
-2	118	3.6%
-1	178	5.4%
0	314	9.6%
1	322	9.8%
2	325	9.9%
3	268	8.2%
4	215	6.6%
5	164	5.0%
6	144	4.4%
7	111	3.4%
8	109	3.3%
9	74	2.3%
>=10	276	8.4%
Total	**1888**	

[a]Includes unvaccinated cases with value 'Not Applicable' in the Immunised field, eg. those tested prior to vaccinations becoming available in New Zealand or individuals aged under 16 who were ineligible for vaccination.

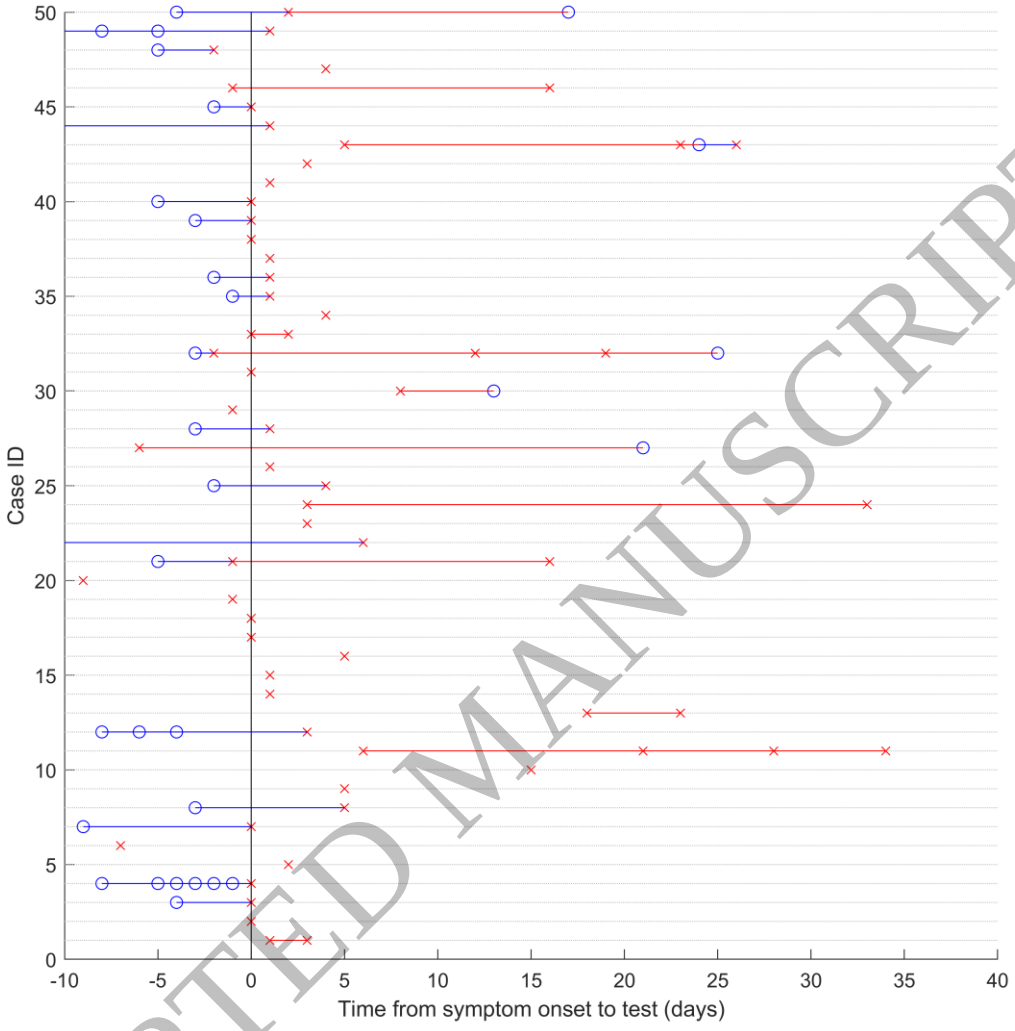
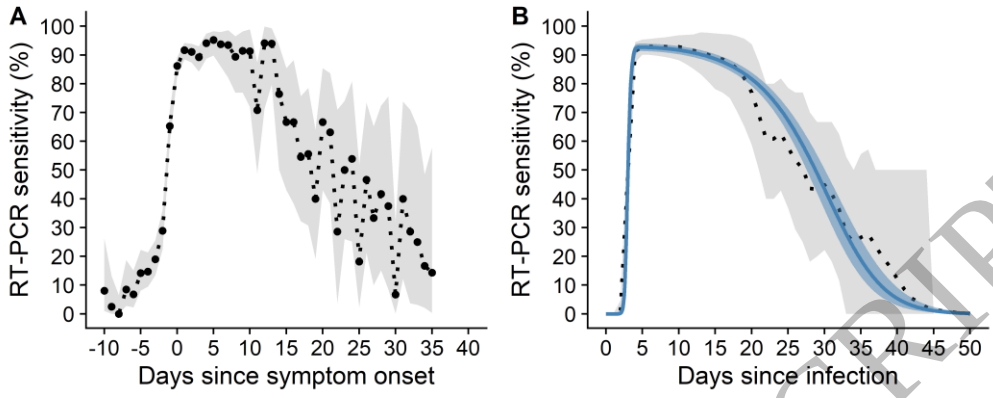


Figure 1
159x153 mm (.02 x DPI)



1

2 **Figure 2**
3 **159x86 mm (.02 x DPI)**

4

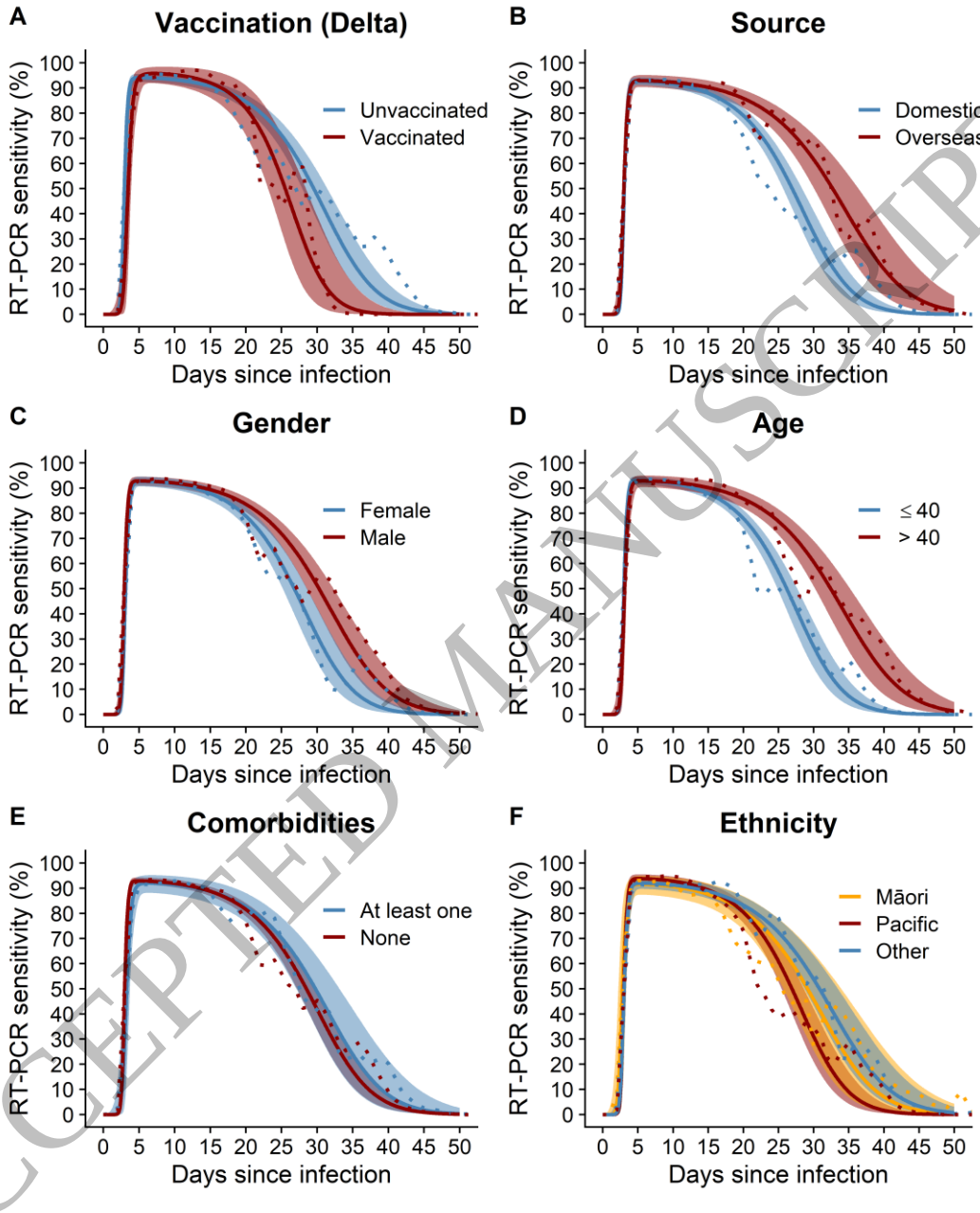


Figure 3
159x212 mm (.02 x DPI)