

SUPPLEMENTARY MATERIALS

Supplementary Methods

Deprivation was assessed using the index of multiple deprivation (IMD), a score based on lower layer super output areas with average population of 1500 people and incorporating seven domains to produce an overall relative measure of deprivation (income, employment, education, skills and training, health and disability, crime, barriers to housing services and living environment) (<https://www.gov.uk/government/statistics/english-indices-of-deprivation-2019>).

Blood was not taken at any visit where any person in the household had classic COVID-19 symptoms (fever, cough or anosmia). Blood was analysed at the University of Oxford. Antibody titres were considered positive above 8 million units[1], and equivocal between 4 and 8 million units[2].

As the Ct distribution was skewed to the left, we assessed independent predictors using median (quantile) regression. Results were broadly similar using random effects model for mean Ct with a random effect per household. We used 5 knot natural cubic splines (knots at the 10th/25th/50th/75th/90th percentiles of observed unique values) to assess non-linearity in the effect of calendar time, age and deprivation (index of multiple deprivation rank). Multivariable models for Ct values were constructed by first choosing the more strongly univariably predictive factor from the collinear variables (symptoms at/around the test, number of genes detected/supporting evidence for each positive) then using backwards elimination on the remaining variables. All analyses were conducted in Stata 14.1.

REFERENCES

1. National Sars-CoV-Serology Assay Evaluation Group. Performance characteristics of five immunoassays for SARS-CoV-2: a head-to-head benchmark comparison. *Lancet Infect Dis* **2020**.
2. Eyre DW, Lumley SF, O'Donnell D, et al. Differential occupational risks to healthcare workers from SARS-CoV-2 observed during a prospective observational study. *Elife* **2020**; 9.

Supplementary Table 1 Association between characteristics and Ct values

(A) Univariable effects and main model considering all factors

| Factor | Levels | n (N=1892) | Median (IQR) Ct per sample unadjusted | p (vs reference) | Adjusted difference in median Ct vs reference† (95% CI) | p (vs reference) |
|--|------------------|---------------|---|---------------------|---|---------------------|
| Independent predictors | | | | | | |
| Overall (at reference category) | | - | - | - | 33.7 (32.9 to 34.5) | |
| Number of genes detected | 1 | 345 | 33.6 (32.3-34.6) | - | - | - |
| | 2 | 185 | 31.5 (29.8-32.8) | 0.004 | -2.5 (-3.8 to -1.2) | 0.0002 |
| | 3 | 1362 | 22.8 (18.2-27.4) | <0.0001 | -10.4 (-11.3 to -9.5) | <0.0001 |
| Any evidence of symptoms around the positive test* | no | 1010 | 28.5 (20.8-32.5) | - | - | - |
| | yes (any) | 866 | 24.2 (19.1-29.1) | <0.0001 | -1.7 (-2.4 to -1.0) | <0.0001 |
| | yes (classic**) | 590 | 22.9 (18.9-27.7) | <0.0001 | -2.1 (-3.0 to -1.3) | <0.0001 |
| | yes (other only) | 276 | 26.6 (19.8-31.1) | 0.01 | -1.1 (-2.1 to -0.0) | 0.04 |
| | missing | 16 | 24.8 (17.4-27.8) | | (excluded) | |
| Any prior positive in the study | no | 1516 | 25.7 (19.1-31.4) | 0.07 | -2.8 (-3.7 to -2.0) | <0.0001 |
| | yes | 376 | 27.1 (22.7-30.7) | - | - | - |
| Collinear factors with similar effects to those above | | | | | | |
| Any evidence of symptoms at the positive test | no | 1214 | 27.9 (20.1-32.2) | - | - | - |
| | yes (any) | 589 | 23.9 (19.1-28.3) | <0.0001 | -1.1 (-2.0 to -0.3) | 0.008 |
| | yes (classic**) | 404 | 23.1 (19.2-27.0) | <0.0001 | -1.2 (-2.2 to -0.2) | 0.02 |
| | yes (other only) | 185 | 25.5 (18.9-30.2) | 0.009 | -1.1 (-2.4 to +0.2) | 0.08 |
| | missing | 89 | 27.8 (22.9-31.8) | | (excluded) | |
| No evidence of independent effect† | | | | | | |
| First test in study (ie positive at first study test) | no | 1267 | 26.0 (19.6-30.9) | 0.19 | | 0.20 |
| | yes | 625 | 26.7 (19.8-31.4) | - | | - |
| Sex | male | 975 | 25.9 (19.4-31.1) | - | | - |
| | female | 917 | 26.4 (20.0-31.1) | 0.29 | | 0.44 |
| Ethnicity | white | 1679 | 26.0 (19.5-31.1) | - | | - |
| | non-white | 201 | 27.3 (21.3-31.3) | 0.11 | | 0.71 |

| Factor | Levels | n (N=1892) | Median (IQR) Ct per sample unadjusted | p (vs reference) | Adjusted difference in median Ct vs reference† (95% CI) | p (vs reference) |
|--|--------------------|---------------|---|-----------------------|---|----------------------|
| | missing | 12 | 26.2 (20.3-30.0) | | | |
| Age group (years/school years (SY))‡ | 2-SY6 | 109 | 26.6 (22.0-31.1) | 0.07 | | 0.48 |
| | SY7-SY11 | 135 | 25.8 (19.8-30.7) | 0.21 | | 0.57 |
| | Y12-24 | 354 | 24.3 (19.3-29.4) | - (global p=0.04) | | - (global p=0.63) |
| | 25-34 | 224 | 26.3 (19.7-30.7) | 0.06 | | 0.67 |
| | 35-49 | 354 | 27.3 (20.1-31.8) | 0.001 | | 0.78 |
| | 50-69 | 504 | 26.8 (19.9-31.9) | 0.002 | | 0.83 |
| | 70+ | 212 | 26.1 (18.1-32.4) | 0.10 | | 0.20 |
| Deprivation quintile †† | 1 (most deprived) | 358 | 24.4 (18.9-29.9) | 0.001 | | 0.04 |
| | 2 | 364 | 25.5 (20.2-30.9) | 0.04 | | 0.75 |
| | 3 | 369 | 26.6 (19.8-30.9) | 0.55 | | 0.31 |
| | 4 | 373 | 26.9 (20.0-31.9) | 0.89 | | 0.71 |
| | 5 (least deprived) | 428 | 27.1 (19.9-31.9) | - (global p=0.004) | | - (global p=0.27) |
| Factor reflecting potential consequence | | | | | | |
| Anyone else in household ever positive | no | 1150 | 28.1 (27.5-28.7) | - | | - |
| | yes | 742 | 23.4 (22.6-24.1) | <0.0001 | | - |

* any evidence of symptoms at the visit where the swab was positive or visits either side of this, regardless of time.

** cough, fever or loss of taste or smell.

† adjusted column shows the effect of adding this variable into a model including number of genes detected, evidence of symptoms around the test and whether the participant had any prior positive in the study (1876 complete cases). Results similar including symptoms around the test instead (1803 complete cases).

‡ linear effect of age on Ct p=0.06 univariable, p=0.59 multivariable; non-linear 0.08 and 0.92 respectively.

†† linear effect of deprivation score on Ct p=0.002 univariable, p=0.30 multivariable; non-linear 0.01 and 0.25 respectively.

Note: p-values (vs the reference category) from univariable or multivariable median regression. Main multivariable regression models include number of genes detected, evidence of symptoms around the test and whether the participant had any prior positive in the study (complete cases); similar (lack of) effect of factors shown above adjusting for symptoms and previous positives only. All multivariable associations similar adjusting for calendar date of visit using a 5 knot natural cubic spline; univariable associations slightly attenuated.

(B) Multivariable model excluding potential mediators of effects of demographics

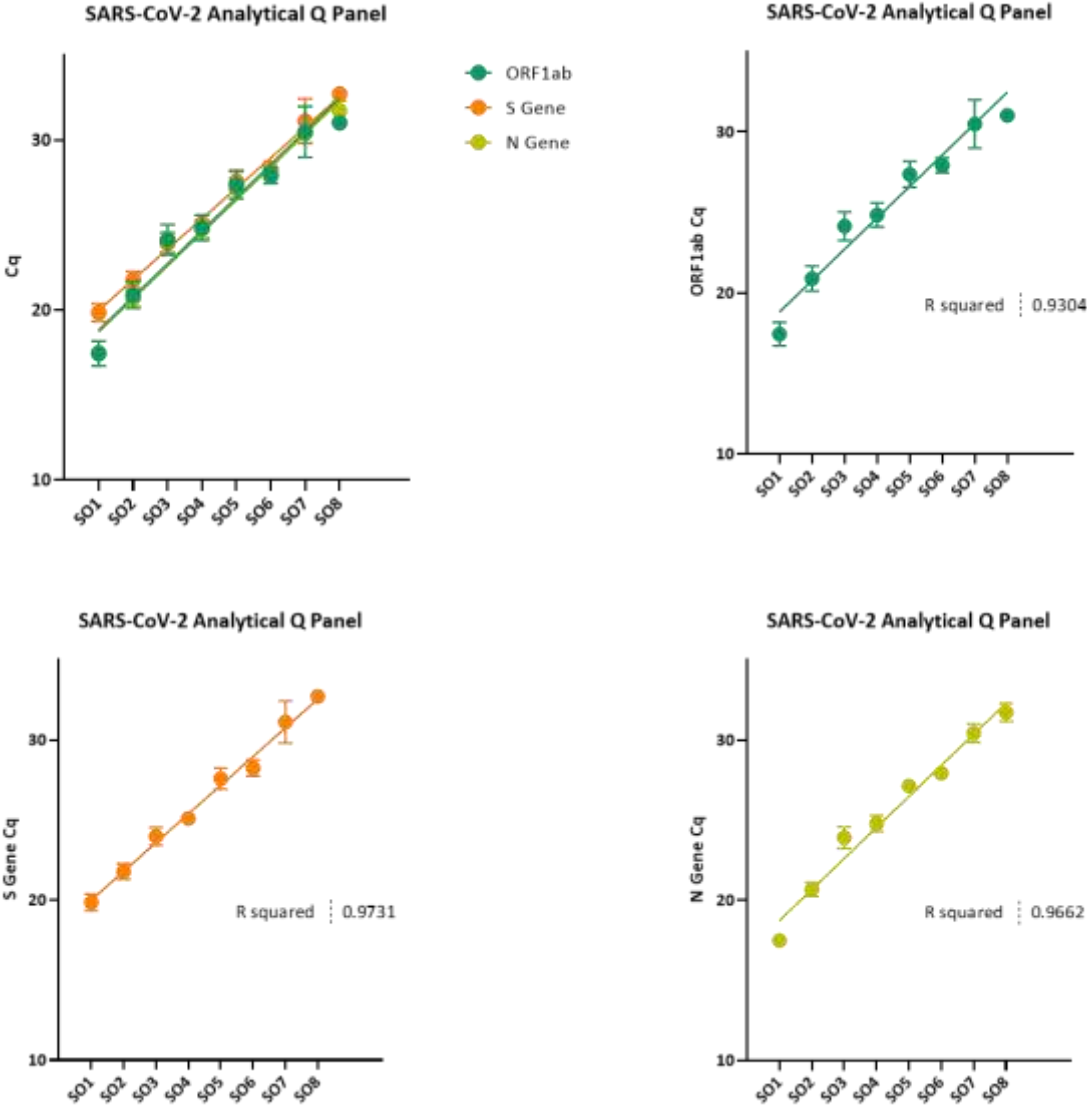
| Factor | Levels | Adjusted difference in median Ct vs reference† (95% CI) | p (vs reference) |
|---|--|---|-------------------|
| Overall (at reference category) | | 32.7 (29.4 to 35.9) | - |
| Any prior positive in the study | no (ie first positive per-participant) | -2.4 (-3.7 to -1.0) | 0.0005 |
| | yes | - | - |
| First test in study (i.e. positive at first study test) | no | -1.4 (-2.5 to -0.2) | 0.02 |
| | yes | - | - |
| Sex | male | - | - |
| | female | +0.5 (-0.5 to +1.5) | 0.31 |
| Ethnicity | white | - | - |
| | non-white | +1.2 (-0.4 to +2.9) | 0.14 |
| Age group (years/school years (SY))‡ | 2-SY6 | +1.8 (-0.6 to +4.2) | 0.14 |
| | SY7-SY11 | +1.2 (-1.0 to +3.4) | 0.28 |
| | Y12-24 | - | - (global p=0.24) |
| | 25-34 | +1.3 (-0.5 to +3.2) | 0.16 |
| | 35-49 | +2.1 (+0.5 to +3.8) | 0.01 |
| | 50-69 | +1.7 (+0.1 to +3.2) | 0.03 |
| | 70+ | +0.7 (-1.2 to +2.6) | 0.46 |
| Deprivation quintile †† | 1 (most deprived) | -2.6 (-4.2 to -1.1) | 0.001 |
| | 2 | -0.9 (-2.5 to +0.7) | 0.27 |
| | 3 | -0.5 (-2.1 to +1.0) | 0.52 |
| | 4 | -0.3 (-1.8 to +1.2) | 0.70 |
| | 5 (least deprived) | - | - (global p=0.01) |

† also adjusted for calendar date using 5 knot natural cubic spline. Reference date 18 July 2020.

‡ linear effect of age on Ct p=0.31; non-linear p=0.14.

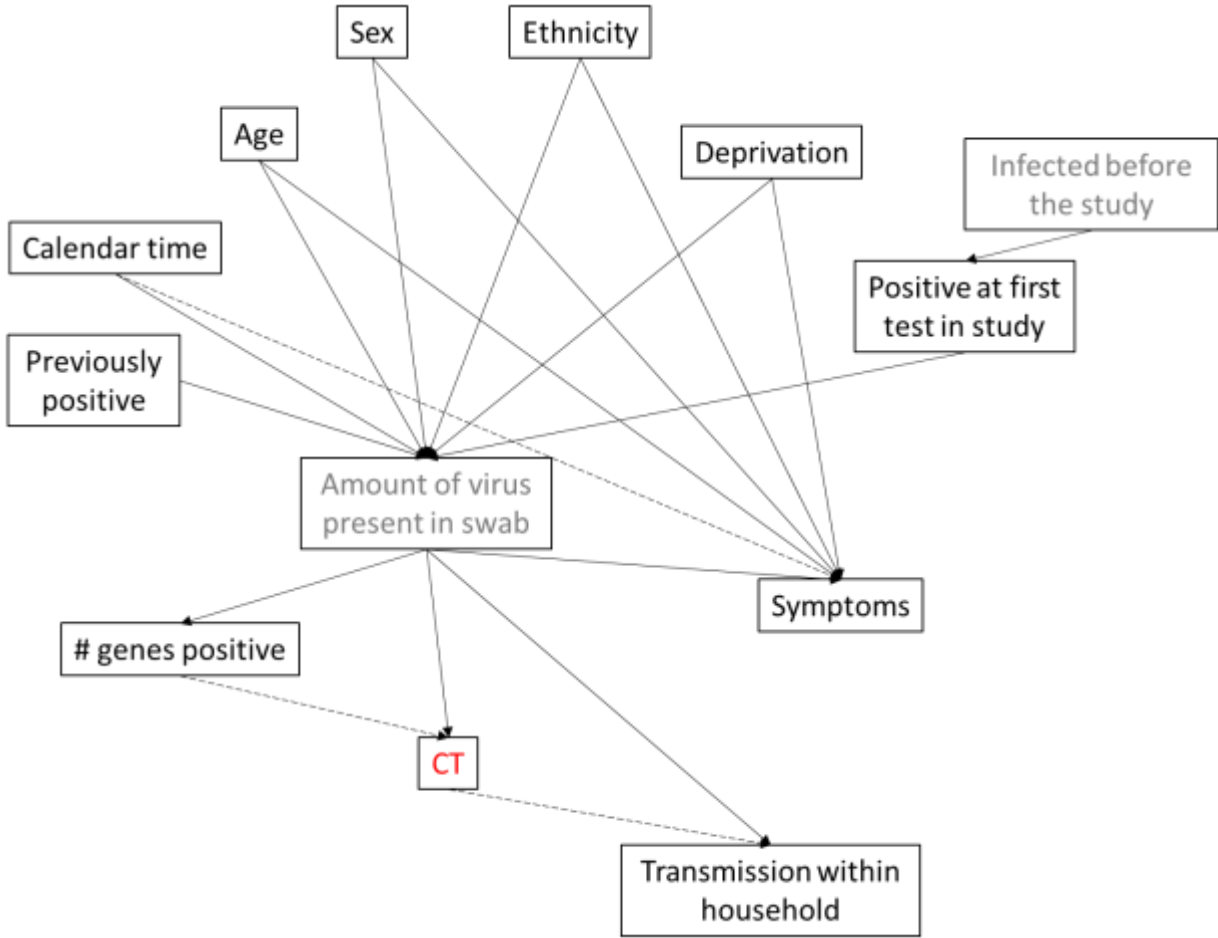
†† linear effect of deprivation score on Ct p=0.001 (+0.6 (+0.2 to +0.9) higher Ct per quintile higher; non-linear p=0.02.

Supplementary Figure 1 Relationship between Ct values and viral load



Note: Mean (+/- 95% CI) CT value plotted against the Qnostics linearity panel (SCV2AQP01-A) with target concentrations from 0 – 1000000 dC/ml (S01=1,000,000 dC/ml, S02=100,000 dC/ml, S03=10,000 dC/ml, S04=5,000 dC/ml, S05=1,000 dC/ml, S06=500 dC/ml, S07=100 dC/ml, S08=50 dC/ml)

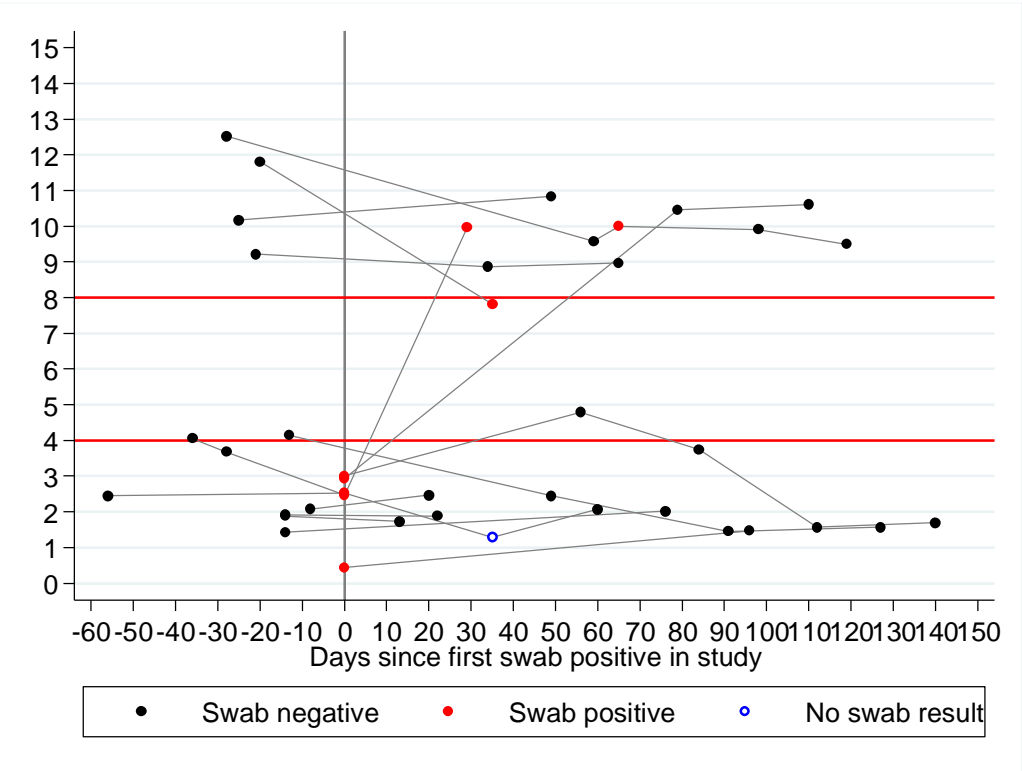
Supplementary Figure 2 Directed acyclic graph of potential relationships between factors



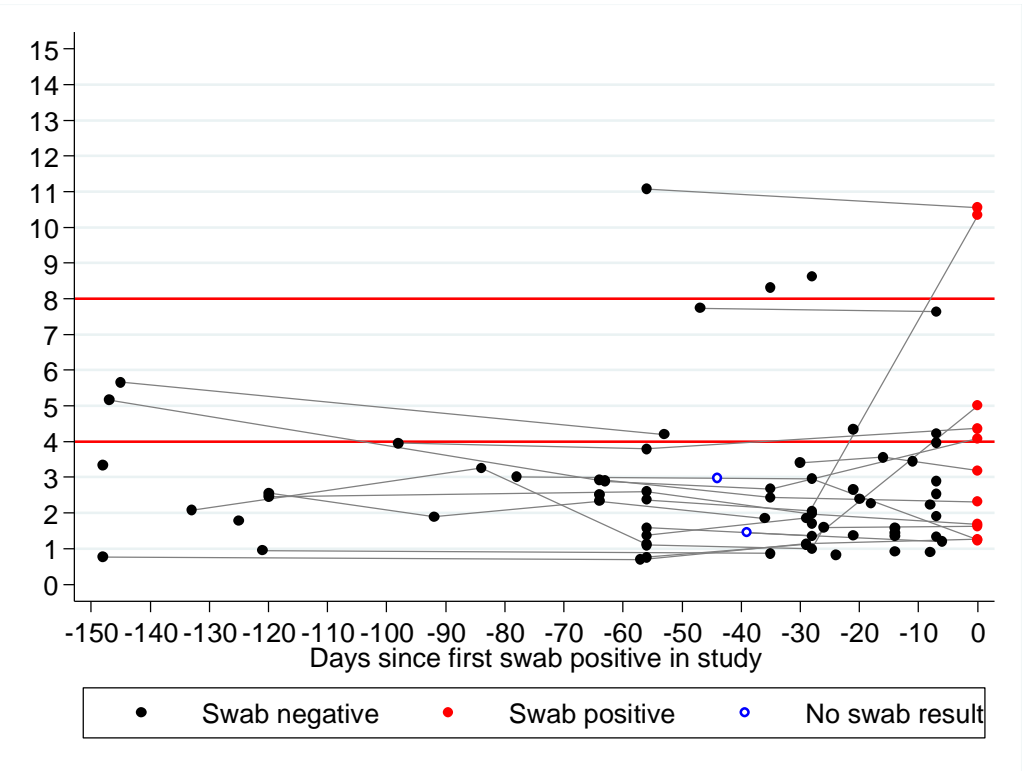
Note: unobserved factors shown in gray, outcome in red.

Supplementary Figure 3 Antibody trajectories in participants with results both before and after their first positive swab (A), only up to and including their first positive swab only (B), and at or after their first positive swab only (C)

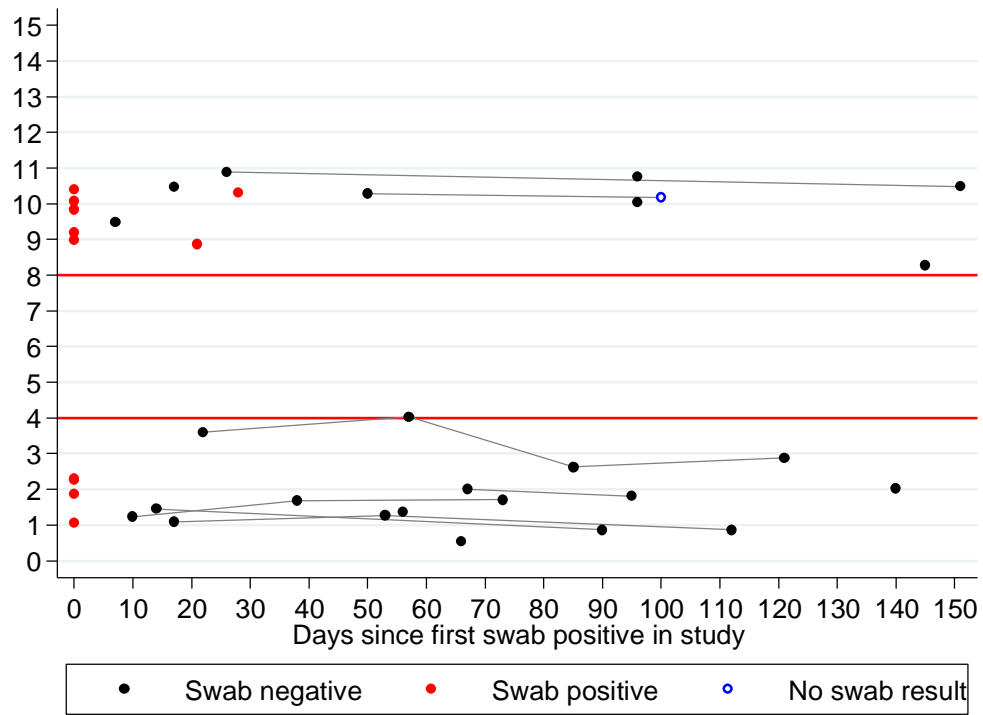
A



B



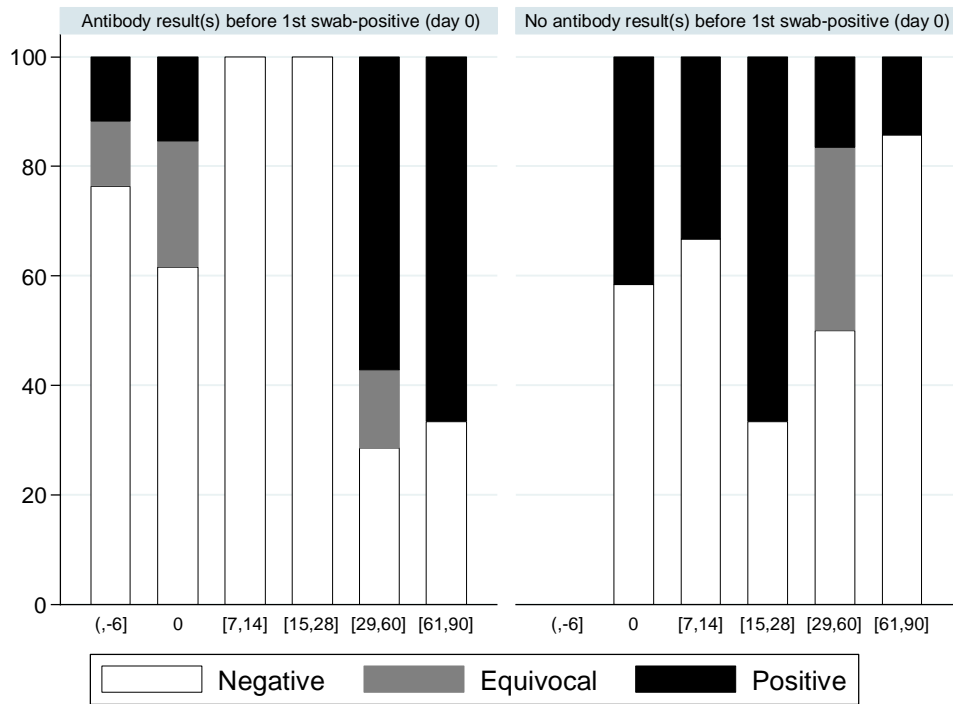
C



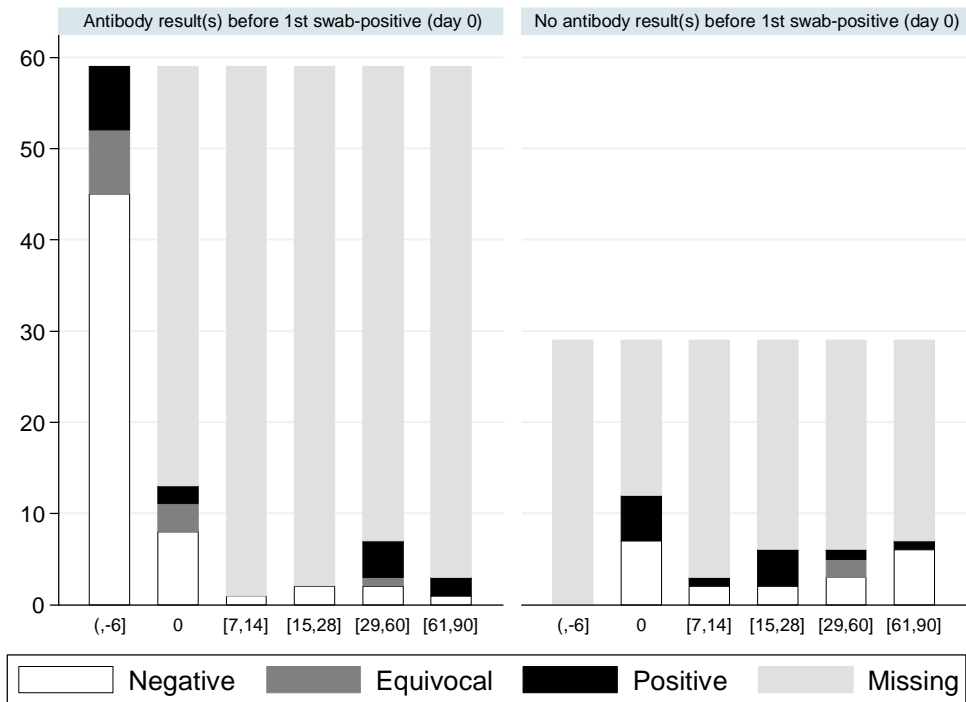
Note: lines join measurements from the same individual. T=0 is first positive swab in the study – blood was not always drawn at this visit; for example blood draws were not done if the participant had symptoms.

Supplementary Figure 4 Percentage (A) and number (B) with antibody responses over time

A



B



Supplementary Figure 5 Individual antibody response trajectories

