**Supplemental Appendix**

# Aerosol and surface stability of SARS-2 (HCoV-19) compared to SARS-CoV

Neeltje van Doremalen1\*, Trenton Bushmaker1\*, Dylan H. Morris2\*, Myndi G. Holbrook1, Amandine Gamble3, Brandi N. Williamson1, Natalie J. Thornburg4, Susan I. Gerber4, James O. Lloyd-Smith3,5, Emmie de Wit1, Vincent J. Munster1

1. Laboratory of Virology, Division of Intramural Research, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Hamilton, MT, USA
2. Dept. of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ, USA
3. Dept. of Ecology and Evolutionary Biology, University of California, Los Angeles, Los Angeles, CA, USA
4. Division of Viral Diseases, National Center for Immunization and Respiratory Diseases, Centers for Disease Control and Prevention, Atlanta, GA, USA.
5. Fogarty International Center, National Institutes of Health, Bethesda, MD, USA

\* These authors contributed equally to this article

# Supplemental methods

Bayesian regression model description

Here we describe in detail the Bayesian regression model that we used to estimate exponential decay rates and half-lives.

In the model notation that follows, the symbol ~ denotes that a random variable is distributed according to the given distribution. Normal distributions are parametrized as Normal(mean, standard deviation). Positive-constrained normal distributions (“Half-Normal”) are parametrized as Half-Normal(mode, standard deviation). We use <Distribution Name>CDF(x, parameters) to denote the cumulative distribution function of a probability distribution, so for example NormalCDF(5, 0, 1) is the value of the Normal(0, 1) cumulative distribution function at 5.

Our data consist of 10 experimental conditions: 2 viruses (HCoV-19 and SARS-CoV) by 5 environmental conditions (aerosols, polypropylene, stainless steel copper and cardboard). Each has three replicates, and multiple time-points for each replicate. We analyze the two viruses separately. For each, we denote by yijk the measured log10 titer in experimental condition *i* during replicate *j* at time-point *k*. To construct our likelihood function, we need to know the probability of observing a given log10 titer measurement yijk given values of the parameters.

Because our titer data are estimated and recorded in increments of 1/nwells log10TCID50/mL, where nwellsis the number of wells used for endpoint titration, our log10 titer values are interval-censored – only known to within a range of width 1/nwells. In addition, there is a degree of measurement noise in the titration process itself.

To model this, we assume that in each experimental condition *i*, there is a true underlying log10 titer xijk that is measured with some amount of experimental noise or error εijk and then observed as an interval-censored value yijk ≈ xijk + εijk. We model the measurement errors εijk as Normally distributed with a standard deviation σi that is shared by all samples in the given experimental condition; this reflects the fact that some experimental setups may be more or less noisy than others.

εijk ~ Normal(0, σi)

We model the probability of observing an interval-censored log10 titer value yijk given a true underlying log10 titer xijk and a measurement error standard deviation σi as:

P(yijk | xijk, σi ) = NormalCDF(yijk, xijk, σi) – NormalCDF(yijk – 1/nwells, xijk, σi)

This reflects the probability given a true value xijk plus the measurement error xijk + εijk falls between yijk – 1/nwells and yijk. Due to the log10 titer imputation technique used, a titer in that range is most likely to be rounded up and reported as yijk.

The detection limit of our experiment is 0.5 log10 TCID50/mL. The probability of observing an undetectable measured log10 titer value yijk given a true log10 titer value xijk is given by:

P(yijk ≤ 0.5 | xijk, σi) = NormalCDF(0.5, xijk, σi)

We then model each replicate *j* for experimental condition *i* as starting with some true initial log10 titer xij(0) = xij0. We assume that viruses in experimental condition *i* decay exponentially at a rate λi over time *t*. It follows that

xij(t) = xij0 – λit

where tk is the kth measured time-point.

Model prior distributions

We place a weakly informative Normal prior distribution on the initial log10 titers xij0 to rule out implausibly large or small values (e.g. in this case undetectable log10 titers or log10 titers much higher than the deposited concentration), while allowing the data to determine estimates within plausible ranges:

xij0 ~ Normal(4.5, 2.5)

We likewise placed a weakly informative Half-Normal prior on the exponential decay rates λi:

λi ~ Half-Normal(0.5, 4)

We placed a weakly informative Half-Normal prior on the standard deviations of the experimental error distributions σi:

σi ~ Half-Normal(0, 2)

Markov Chain Monte Carlo Methods

We drew posterior samples using Stan, which implements a No-U-Turn Sampler (a form of Markov Chain Monte Carlo). We ran four replicate chains from random initial conditions for 2000 iterations, with the first 1000 iterations as a warmup/adaptation period. We saved the final 1000 iterations from each chain, giving us a total of 4000 posterior samples. We assessed convergence by inspecting trace plots and examining *R̂* and effective sample size (neff) statistics (*R̂* for all parameters ≤ 1.003, neff for all parameters ≥28% of total samples).

# Supplemental Figures

Figures S1–S5 show Bayesian fits to individual replicate virus decay data for each virus. Replicates are shown in panel columns, viruses in panel rows. Lines are 50 random draws per panel from the posterior distribution of fitted lines, to show level of uncertainty. Time axis on same scale across all figures, shown out to the latest time taken to reach an undetectable titer (96 hours, for all replicates of both viruses on plastic).

A close up of a map

Description automatically generated

Figure S1. Individual replicate fits for aerosols. Columns show replicates, rows show virus (HCoV-19 above, SARS-CoV-1 below). Lines are 50 random draws per panel from the posterior distribution of fitted lines, to show level of uncertainty.

A picture containing text, sky, map, indoor

Description automatically generated

Figure S2. Individual replicate fits for plastic (polypropylene). Columns show replicates, rows show virus (HCoV-19 above, SARS-CoV-1 below). Lines are 50 random draws per panel from the posterior distribution of fitted lines, to show level of uncertainty.

A picture containing text, indoor, sky, map

Description automatically generated

Figure S3. Individual replicate fits for steel. Columns show replicates, rows show virus (HCoV-19 above, SARS-CoV-1 below). Lines are 50 random draws per panel from the posterior distribution of fitted lines, to show level of uncertainty.

A close up of a map

Description automatically generated

Figure S4. Individual replicate fits for copper. Columns show replicates, rows show virus (HCoV-19 above, SARS-CoV-1 below). Lines are 50 random draws per panel from the posterior distribution of fitted lines, to show level of uncertainty.

A close up of a map

Description automatically generated

Figure S5. Individual replicate fits for cardboard. Columns show replicates, rows show virus (HCoV-19 above, SARS-CoV-1 below). Lines are 50 random draws per panel from the posterior distribution of fitted lines, to show level of uncertainty. Fits are substantially poorer for SARS than for HCoV-19, and data do not follow a linear downward trend over time, suggesting that the difference in observed decay rates should be interpreted with caution.