Supplementary material

A Time course of virus shedding

Suppose virus is accumulated at an infection site and released into the intestinal tract with concentration $c_1(t)$, where it is transported with peristalsis, to be excreted ultimately. The concentration $c_2(t)$ of virus leaving the intestines then can be described as resulting from

$$\begin{cases} c_1'(t) = -\alpha c_1(t); & c_1(0) = A \\ c_2'(t) = +\alpha c_1(t) - \beta c_2(t); & c_2(0) = 0 \end{cases} \tag{A.1}$$

where A is the initial concentration of virus at the primary infection site, α and β are constants defined by the transport rate and effective volumes of the compartments within the intestinal tract [32]. The solution

$$\begin{cases} c_1(t) = A\mathrm{e}^{-\alpha t} \\ c_2(t) = \frac{\alpha}{\beta-\alpha} A\mathrm{e}^{-\alpha t}\left(1 - \mathrm{e}^{-(\beta-\alpha)t}\right) \end{cases} \tag{A.2}$$

The observed virus concentration can then be written as

$$C(t|\alpha,\beta) = C_0\mathrm{e}^{-\alpha t}(1 - \mathrm{e}^{-(\beta-\alpha)t}) \tag{A.3}$$

If the constant C_0 is defined as

$$C_0 = \frac{\alpha}{\beta-\alpha}\left(\frac{\alpha}{\beta}\right)^{\frac{\beta}{\alpha-\beta}} \tag{A.4}$$

then the peak virus concentration is 1. When virus titres are measured on an inverted log scale ($-\log(C) = u$) the regression model is

$$u(t|\alpha,\beta,c,d) = c - d\log(C(t|\alpha,\beta)) \tag{A.5}$$

B Model fitting

The above model can be fitted to the observed Ct values by assuming they have a normally distributed error with expected value u and standard deviation σ (or precision $\tau = 1/\sigma^2$, in a mixed model framework with distributions for α and β and c and d, describing their (joint) variation in the observed population

$$(\alpha, \beta, c, d) \sim N(\boldsymbol{\mu}, \boldsymbol{\Omega}) \tag{A.6}$$

with mean vector $\boldsymbol{\mu} = (\mu_\alpha, \mu_\beta, \mu_c, \mu_d)$ and precision matrix $\boldsymbol{\Omega}$. Note that this setup can be easily adapted to account for censoring (Ct $\geq$ 40).

For asymptomatic subjects the onset of symptoms (onset of virus shedding) is missing. Therefore it is inconvenient to take onset of shedding as the origin ($t = 0$) of the shedding model. Instead, the time from the first sample (Δt), is used

$$u(t + \Delta t|\alpha, \beta, c, d) = c - d\log(C(t + \Delta t|\alpha, \beta)) \tag{A.7}$$

For symptomatic cases Δt is assumed known: the time between symptom onset and first faecal sample. For asymptomatic cases where Ct data are present, Δt can be estimated.

Using the date of the first faecal sample as a reference, estimates of the period Δt_{pred} were obtained: the time between onset of shedding and the first sample. In symptomatic cases with faecal samples present there is good agreement between Δt_{pred} and observed times between the onset of symptoms and the first sample, indicating that symptoms tend to appear within a day after the start of virus shedding. In asymptomatic subjects the onset of symptoms is missing, and the fitting procedure thus produces estimates of the onset of shedding, based on observed virus shedding patterns [17].

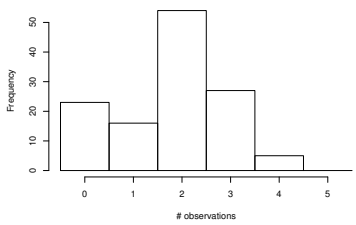
The source code for the JAGS model:

```
model{
  for(subj in 1:n.subj) {
    a0[subj] <- b0[subj]/(exp(u0[subj]*b0[subj])-1)
    d0[subj] <- (exp(u0[subj]*b0[subj])-1)*v0[subj]/b0[subj]
```

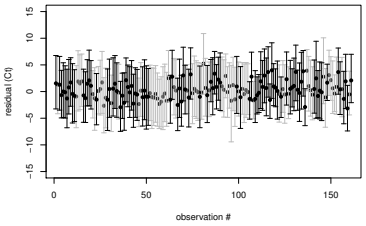
```
    fac[subj] <- ((a0[subj]+b0[subj])/b0[subj])*
                 ((a0[subj]+b0[subj])/a0[subj])^(a0[subj]/b0[subj])
    offs.sympt[subj] ~ dnorm(mu.offs,tau.offs)
    loglat[subj] ~ dnorm(mu.loglat,tau.loglat)
    offs.shed[subj] <- offs.sympt[subj]-exp(loglat[subj])
    for(obs in 1:n.obs[subj]) {
      t[subj,obs] <- t.obs[subj,obs]-offs.shed[subj]
      vir[subj,obs] <- ifelse(t[subj,obs] <= 0, 60,
                       c0[subj]-d0[subj]*log(
                       exp(-a0[subj]*t[subj,obs])*
                       (1-exp(-b0[subj]*t[subj,obs]))
                       *fac[subj]))
      ct.cens[subj,obs] ~ dinterval(ct.obs[subj,obs],censorlimit)
      ct.obs[subj,obs] ~ dnorm(vir[subj,obs],tau.obs)
    }
    u0[subj] <- exp(par[subj,1])
    b0[subj] <- exp(par[subj,2])
    c0[subj] <- exp(par[subj,3])
    v0[subj] <- exp(par[subj,4])
    par[subj,1:4] ~ dmnorm(mu.par,tau.par)
  }
  mu.par ~ dmnorm(mu.hyp,tau.hyp)
  tau.par ~ dwish(omega,4)
  tau.obs ~ dgamma(tau.obs.hyp[1],tau.obs.hyp[2])
}
```

A multivariate normal hyperprior was used for the means $\boldsymbol{\mu}$ (`mu.par`), with mean vector (2.0,-3.0,2.7,-0.7) and a diagonal matrix $0.0001\boldsymbol{I}_4$ for its precision (`tau.par`). For the precision matrix $\boldsymbol{\Omega}$ (`omega`) an informed Wishart prior ($\boldsymbol{I}_4$) was used. The prior for Δt was normal (with mean `mu.offs` = -5 and precision `tau.offs` = 0.01); the log of the latency between onset of shedding and onset of symptoms was modelled normal (mean `mu.loglat` = -5, precision `tau.loglat` = 10). Measurement error in Ct values was incorporated by assuming gamma (4,4) distributed precision (`tau.obs.hyp`).

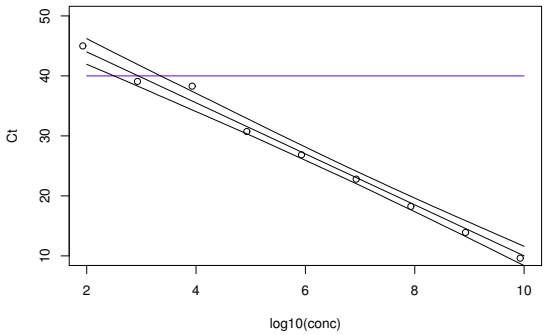
C Additional output



(a) Numbers of observations



(b) Residuals



(c) Standard curve

Figure A1: (a) Numbers of observations (Ct values) per individual subject. Cases without any observations were not included in the analyses. (b) Residuals (Monte Carlo sample of difference between predicted and observed Ct) by observation (any observed Ct in any subject). Mean values and 95% range, symptomatic (black) and asymptomatic (grey) subjects. Observations with Ct $\geq$ 40 are not shown here. (c) Calibration of the quantitative PCR for GII.4 NoV.Observed Ct values for a dilution series of a standard suspension, and linear calibration curve, with 95% predictive intervals. In the regression model for the calibration curve the highest Ct value ($\geq$ 40) is treated as a censored observation.

	Asymptomatic			Symptomatic		
	Time to peak	Peak level	Duration	Time to peak	Peak level	Duration
Peak level	0.481			0.364		
Duration	0.973	0.467		0.873	0.407	
Area under shedding curve	0.668	0.960	0.664	0.557	0.960	0.611

Table A1: Correlation coefficients for the shedding characteristics (means of individual estimates).

C.1 Patients and staff

	Time to peak				Peak level			
	Staff sympt	Patient sympt	Staff asympt	Patient asympt	Staff sympt	Patient sympt	Staff asympt	Patient asympt
Staff, sympt		0.600	0.522	0.586		0.642	0.547	0.640
Patient, sympt				0.482				0.503
Staff, asympt				0.560				0.582
Patient, asympt								

	Duration				Area under shedding curve			
	Staff sympt	Patient sympt	Staff asympt	Patient asympt	Staff sympt	Patient sympt	Staff asympt	Patient asympt
Staff, sympt		0.815	0.579	0.750		0.700	0.570	0.686
Patient, sympt				0.419				0.496
Staff, asympt				0.669				0.624
Patient, asympt								

Table A2: Comparison of shedding characteristics among staff and patients, with or without symptoms. Fraction (of posterior MC sample) with positive difference between categories (columns vs. rows). Symptomatic patients seem to shed longer than asymptomatic patients and staff, but nowhere the fraction with difference > 0 exceeds 0.95.