

Supplementary Materials - Estimating the trend of COVID-19 in Norway by combining multiple surveillance indicators

Supplementary methods

Estimating prevalence from for the “Fraction Positive - Survey” indicator

For the “Fraction Positive - Survey” indicator we want to estimate the symptomatic prevalence of those surveyed, but not everyone with symptoms chose to get tested. We therefore assume that the test positivity rate is equal among all participants who report having cold symptoms and use a Bayesian model to estimate the overall prevalence as follows:

$$\begin{aligned} p_{covid} &= p_{sympt} * p_{pos} \\ N_{sympt} &\sim \text{binomial}(N, p_{sympt}) \\ N_{pos} &\sim \text{binomial}(N_{test}, p_{pos}), \end{aligned}$$

where N is the total number of participants, N_{sympt} the number with symptoms, N_{pos} the number of symptomatic participants who took a test and N_{test} the number who tested positive. Using this setup we can estimate the overall symptomatic prevalence p_{covid} with uncertainty.

Estimating incidence from Prevalence

Incidence, $I(t)$ and prevalence, $P(t)$ are related by:

$$P(t) = \sum_s p(s)I(t-s)$$

where $p(s)$ is the probability of testing positive for someone who was infected at $s = 0$. We use a PCR-test profile as estimated from [1], which is an approximation for the symptometer data which includes both PCR and antigen tests. To estimate the incidence, we then use an approximate Gaussian process model with a Matern Kernel.

$$\text{logit}(I(t)) = i_0 + i(t), i(t) \sim GP(t)$$

Similarly following [2], we estimate incidence from wastewater data. Here, the observed amount of SARS-CoV-2 in the wastewater is a convolution of the incidence and the time-profile of gastrointestinal viral load. We assume a standard deviation of 20% of the mean value for the normalised SARS-CoV-2 concentration. We used a similar Gaussian process model as above. For both approaches we use the implementation presented in [3].

Disaggregation of weekly to daily data

To disaggregate weekly data, we implement a similar model to the models above where daily infections again are given by

$$\text{logit}(I(t)) = i_0 + i(t), i(t) \sim GP(t)$$

We then calculate weekly aggregated data $W_i = \sum_t \text{Ind}(\text{week} = i)I(i)$ where $\text{Ind}(\text{week} = i)$ is an indicator function for the week. These weekly aggregated incidences are then compared to the aggregated data, A

$$A \sim \text{poisson}(W).$$

Using the Stan language we can then get samples for $I(t)$ which we then feed into the regression model to ensure we incorporate the uncertainty from the disaggregation.

Combining estimated growth rates

We combine the estimated growth rates from the different surveillance systems using methods from the field of meta-analysis where we consider the mean, $r_i(t)$ and the standard deviations $\sigma_i(t)$ calculated from the estimated growth rates from each indicator. We implement a random effects meta-analysis model with a random walk for smoothing over time.

$$\begin{aligned} r_i(t) &\sim \text{normal}(\mu_i(t), \sigma_i) \\ \mu_i(t) &\sim \text{normal}(r(t), 0.1) \\ r(t=1) &\sim \text{normal}(0, 0.1) \\ r(t) &\sim \text{normal}(\alpha + \beta r(t-1), \tau) \\ \tau &\sim \text{normal}(0, 0.5) \\ \alpha &\sim \text{normal}(0, 5) \\ \beta &\sim \text{normal}(0, 5) \end{aligned}$$

The usefulness of the combined estimate from all the indicators will depend on the heterogeneity of the individual indicators. We estimate the I^2 indicator for the heterogeneity [4] using the following estimate:

$$I^2(t) = \frac{\tau^2(t)}{\tau^2(t) + s^2(t)}$$

where $s^2(t)$ is the geometric mean of the variances from the growth rates of the individual indicators. All the models are implemented in the Stan probabilistic programming language.

Estimating the Effective Reproduction Number from the Growth Rate

From the growth rates one can estimate reproduction numbers [5] if we assume a gamma distributed generation with shape, α and rate β .

$$R = \left(1 + \frac{r}{\beta}\right)^\alpha.$$

Following the results in [6] we update the parameters in the gamma distribution by the dominating variant in Norway. For each variant (Wuhan, Alpha, Delta and Omicron) we use a mean given by [6]. To be able to estimate the Reproduction Number we then also need the spread of the distribution, here characterised by the coefficient of variation, $cv = \sigma/\mu$. For the ancestral Wuhan variant, we use a coefficient of variation of 1.2 [7], and then 0.73 for Alpha, 0.7 for Delta [8] and 0.75 for Omicron [9]. We model the transition between different variants with a logistic function matched to the data on Variants from ECDC [10]. The transition is then given by:

$$f(t) = \frac{1}{1 + \exp\left(\frac{t-d}{w}\right)},$$

where f is the fraction of infections with the new variant, d is the date when the variant reached 50% and w is the width of the transition. d and w were manually extracted from the data. The value of the mean or the coefficient of variation X_i is then given by:

$$X_i(t) = fX_i^n + (1 - f)X_i^p,$$

where X_i^n is the value for the new variant and X_i^p is the value for the old variant.

Relative incidence

From an estimated growth rate, $r(t)$, we can also estimate an associated relative incidence by

$$I(t) = \prod_{i=1}^t \exp(r(i)).$$

We sample from the estimates of $r(t)$ to quantify the uncertainty in $I(t)$

Supplementary Results

To investigate the effect of changing the amount of smoothing when estimating growth rates, we have estimated growth rates with different window lengths, $2l + 1$, in the negative binomial regression model. In Figure 1 we show the estimated growth rate for the number of reported cases with changing length of the time-window. This clearly illustrates how a longer time window has lower uncertainty and fewer temporal features. Increased smoothing also shrinks the growth rates towards 0.

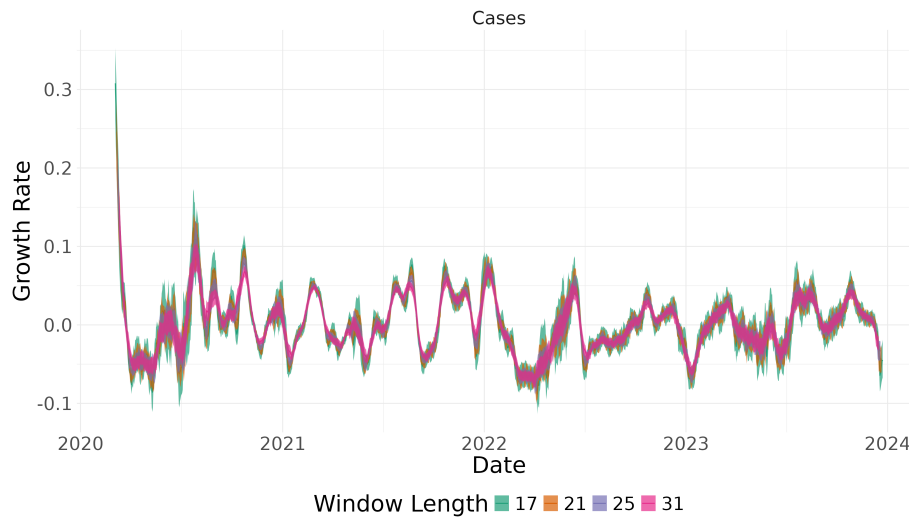


Figure 1: Estimated growth rates for the number of positive cases with varying amount of smoothing controlled by the length of the time-window used, Norway 2020-2023.

Following the same procedures as in the main paper we show in Figure 2, with a window length of 21 days, and Figure 3, with a window length of 31 days, the correlation matrix between the growth rates of the different indicators when we vary the amount of smoothing. Higher smoothing leads to a higher correlation, likely due to two main effects. Since some of our indicators are smoothed by default, for example due to being weekly indicators, it is possible that more smoothing of daily indicators allows us to have more similar amounts of smoothing. Secondly, since smoothing has a tendency to shrink estimates towards 0, one can likely also get some spurious correlations by smoothing too much.

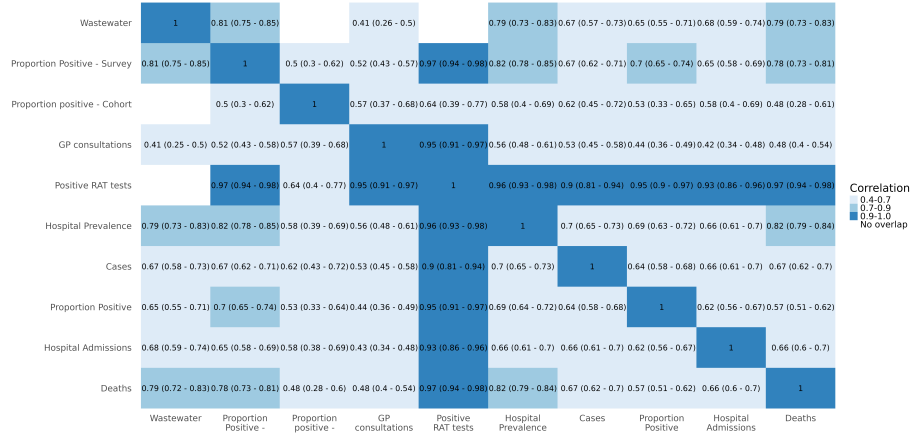


Figure 2: Estimated correlation with 95% confidence intervals between the estimated growth rates of the different surveillance indicators for a smoothing time-window of 21 days, Norway 2020-2023.

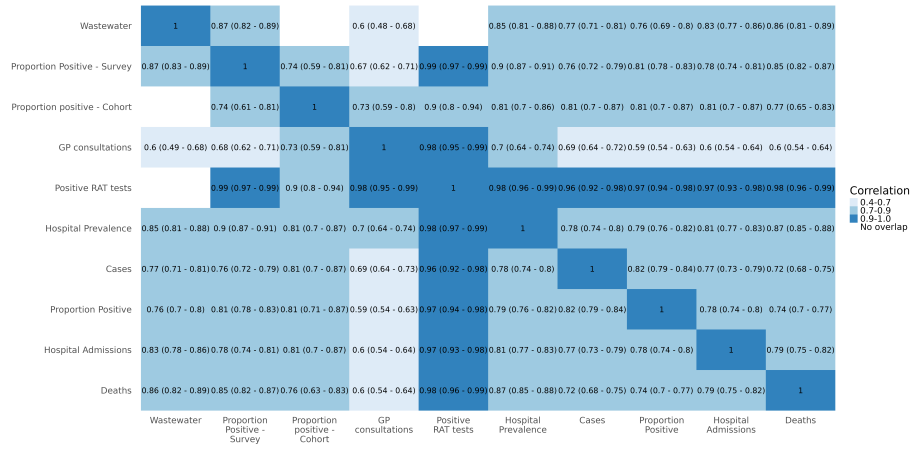


Figure 3: Estimated correlation with 95% confidence intervals between the estimated growth rates of the different surveillance indicators for a smoothing time-window of 31 days, Norway 2020-2023.

References

1. Hellewell J, Russell TW, Matthews R, Severn A, Adam S, Enfield L, et al. Estimating the effectiveness of routine asymptomatic PCR testing at different frequencies for the detection of SARS-CoV-2 infections. *BMC Medicine*. 2021 Apr;19(1):106.
2. Fernandez-Cassi X, Scheidegger A, Bänziger C, Cariti F, Tuñas Corzon A, Ganesanandamoorthy P, et al. Wastewater monitoring outperforms case numbers as a tool to track COVID-19 incidence dynamics when test positivity rates are high. *Water Research*. 2021 Jul;200:117252.
3. Abbott S, Funk S. Estimating epidemiological quantities from repeated cross-sectional prevalence measurements. *medRxiv*; 2022. p. 2022.03.29.22273101.
4. Higgins JPT, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*. 2002 Jun;21(11):1539–58.
5. Wallinga J, Lipsitch M. How generation intervals shape the relationship between growth rates and reproductive numbers. *Proceedings of the Royal Society B: Biological Sciences*. 2006 Nov;
6. Xu X, Wu Y, Kummer AG, Zhao Y, Hu Z, Wang Y, et al. Assessing changes in incubation period, serial interval, and generation time of SARS-CoV-2 variants of concern: A systematic review and meta-analysis. *BMC Medicine*. 2023 Sep;21(1):374.
7. Hart WS, Abbott S, Endo A, Hellewell J, Miller E, Andrews N, et al. Inference of the SARS-CoV-2 generation time using UK household data. Flegg J, Franco E, Kao RR, Conway E, editors. *eLife*. 2022 Feb;11:e70767.
8. Hart WS, Miller E, Andrews NJ, Waight P, Maini PK, Funk S, et al. Generation time of the alpha and delta SARS-CoV-2 variants: An epidemiological analysis. *The Lancet Infectious Diseases*. 2022 May;22(5):603–10.
9. Manica M, Bellis AD, Guzzetta G, Mancuso P, Vicentini M, Venturelli F, et al. Intrinsic generation time of the SARS-CoV-2 Omicron variant: An observational study of household transmission. *The Lancet Regional Health – Europe*. 2022 Aug;19.
10. European Centre for Disease Prevention and Control. European Respiratory Virus Surveillance Summary (ERVISS), 2024, Week 17.