Nowcasting COVID-19 hospitalisations in Germany using reported cases

Stefan Heyder, Thomas Hotz

31.03.2023

Introduction

Judging the severity of the COVID-19 epidemic has been an ongoing challenge since its inception. As immunization against COVID-19 rose, strict enforcement of social distancing rules eased and testing regimes became less strict, case incidences became a less reliable and harder to interpret indicator of epidemic severity. Instead more direct indicators of morbidity, such as the number of deaths and ICU admissions and occupancy have come to the fore. But these indicators are late due to the substantial delays between infection and occurence. An alternative indicator that captures the morbidity caused by COVID-19 but is earlier than the others is the number of hospitalisations of positive COVID-19 cases.

While hospitalisations occur earlier, they still come with substantial delay between the infection and subsequent admission to hospital. Additional difficulties arise due to delays in reporting, i.e. the time it takes until the hospital reports the new case to the national health authorities. The problem of accounting for delays in reporting for occurred, but not yet reported events has been termed **nowcasting**, i.e. forecasting of the indicator at time "now". Predicting the number of hospitalisations is thus a mixture of both forecasting — which reported COVID-19 cases will end up in the hospital — and nowcasting — which cases have yet to be reported — and we will use the term nowcasting in this paper to mean this predictive mixture. We focus on the situation in Germany where data on hospitalisations has been avaiable since April 2021 provided by the German federal health care authorithy, the Robert Koch-Institut (RKI), via Github [19]. In these data the number of hospitalisations is linked to the date of reporting of the associated case, so, in a sense, the term of nowcasting is accurate: we are interested in the "true" value of the indicator today, that will only be observed after a long delay. While this association requires a careful interpretation of the indicator (see the Discussion section of this paper) it was, besides case incidences and ICU occupancy, one of the main official indicators in Germany informing countermeasures and so there is merit in nowcasting it.

The extent of delays is visible in Figure 1: the reported number of hospitalisations will roughly double over the course of twelve weeks. By the aforementioned reporting scheme of hospitalisations there are two reporting dates for a single hospitalised case: the reporting date of the case, i.e. the date when local health authorities were made aware of the positive test, and the reporting date of the hospitalisation, i.e. when the hospitalisation was reported to the RKI. This induces a double weekday effect in the reporting delays which we make visible in Figure 2.

Compared to previous approaches to nowcasting that tend to exclusively focus on modelling the delay distribution with intricate parametric models, our model sidesteps this complex delay structure by decomposing delayed hospitalisations into weekly chunks (Figure 6) and incorporate additional data. As cases and hospitalisations are explicitly linked by the case reporting date we forecast the number of hospitalisations in each chunk based on the current incidences and past fractions of hospitalisations in a comparable weekly chunk. We additionally quantify uncertainty by prediction intervals that are informed by the past performance of our model. This makes our model straightforward to understand, easy to implement and fast to run.

The origin of nowcasting lie in accounting for incurred, but not reported claims in the actuarial sciences [11], delays in reporting for AIDS [25, 12] and other infectious diseases [7].

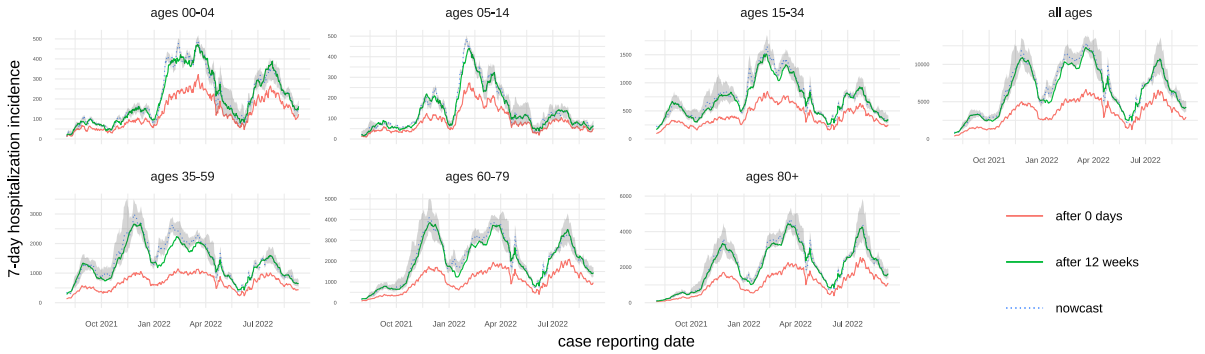

Figure 1: Germany's 7-day hospitalisation incidence changes due to various delays such as time to hospitalisation and delays in reporting. This figure shows the extent of these delays: incidences reported at the present date (red lines) severely underestimate the hospitalisation incidence (green solid lines) that is reported after 3 months. Our nowcasting model (blue dotted lines, 95% prediction intervals in shaded gray) deals with this problem by predicting the hospitalisation incidence based on past cases and their delays to hospitalisation.

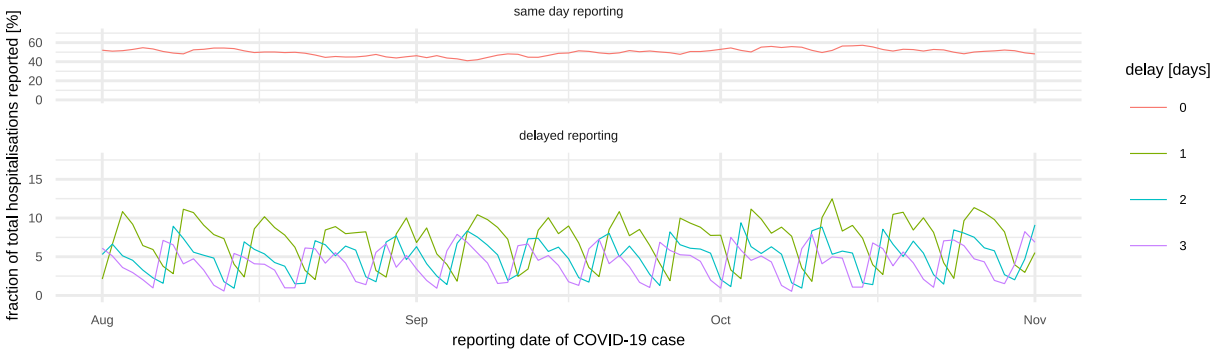

Figure 2: The hospitalisation incidence contains a double weekday effect owed to the reporting of both the COVID-19 case and the subsequent hospitalisation. While the weekday effect of the case reporting date is somewhat mitigated by summing over 7 day periods, the weekday effect of reporting date of the hospitalisation is still present in the data. It is most pronounced for hospitalisations that are reported with delays, i.e. where the case reporting date does not match the reporting date of the hospitalisation.

Nowcasting problems have been analysed with methods from survival analysis [12] and generalized linear regression [25]. In the survial analysis setting one commonly models the reverse time discrete hazard parametrically and assumes multinomial sampling of the final number of cases, potentially accounting for overdispersion. This has been studied with frequentist [14] and bayesian [9, 2] methods. The generalized linear regression approach has origins in the chain ladder model from actuarial sciences [18] and models the observed counts in the reporting triangle by a Poisson or negative binomial distribution.

For both approaches available covariates can be incorporated in a straight-forward way. In the setting of real-time nowcasting it is often beneficial to incorporate epidemic dynamics into the model, this can be done by splines [9, 24] or by a latent process of infections [13].

Nowcasting methods have wide application in accouting for reporting delays [14], early outbreak detection [22, 3], and, in the recent COVID-19 epidemic, improving real-time monitoring of epidemic outbreaks [2, 8, 23, 1].

Evaluating a forecasting model in a real-time public health setting is advantageous as it avoid hindsight bias [5], however nowcasting approach may have difficulties with bias and properly calibrated uncertainty if used in a real-time setting. This includes rapidly changing dynamics [8, 24], both of the delay distribution and the underlying epidemic, retrospective changes in data [14] and long delays with few observed cases [15]. To avoid the aforementioned hindsight bias one can make their predictions publicly available in real-time [17, 4]. For the hospitalisations in Germany we have particiapted in the German COVID-19 NowcastHub [16] since November 2021 where nowcasts are available in a public Github repository [10].

Results

Before evaluating the predictive performance of our model we investigate how the fraction of hospitalisations after one up to four weeks changes over time across different age groups. Figure 3 shows that these fractions are changing slowely over time, especially in the older age groups. Due to smaller numbers of infections and hospitalisations reported in the younger age groups these fractions vary more strongly, occassionally dropping to 0. Across all age groups we observe a steady decline from October 2021 to December 2021 with a steeper drop in fraction of hospitalisations starting with January 2022. The former period corresponds to a time of mandatory testing at the workplace which may improve ascertainment of asymptomatic and less severe cases. The latter effect is most recognizable in the 35-59 age group and coincides with the time that the Omicron variant became dominant in Germany [21]. Additionally there is no visually discernible weekday effect present in Figure 3.

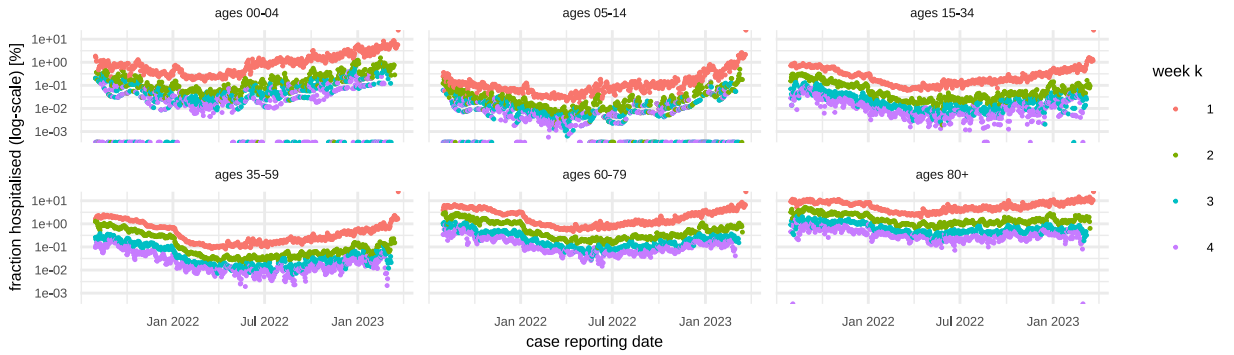

Figure 3: We show the fractions of hospitalisations in the k-th week after case reporting date t of initially reported cases in different age groups, i.e. $p_{t,0,k} = \left(H_{t,7k} - H_{t,7\cdot(k-1)}\right)/I_{t,0}$. Note the log-scale of the y-axis. During periods of low incidence, e.g. July – September, we find large fluctations, but no discernable weekly pattern. With rising case numbers the fractions stabilise and decrease in most age groups. This might be due to changes in testing regime detecting less severe cases. As changes occur on slow time scales, estimating these fractions by Eq. (1) is a promising approach.

In Figure 1 we depict the nowcasts produced from our model including 95% prediction intervals, whose

lengths are based on the past performance of our model. Except for the period from January to April 2022, the model produces reasonable nowcasts with prediction intervals that have sensible widths. In the aforementioned period the nowcasts overpredict the final hospitalisations, except for the oldest age group, and, after a transitionary period, have larger uncertainty.

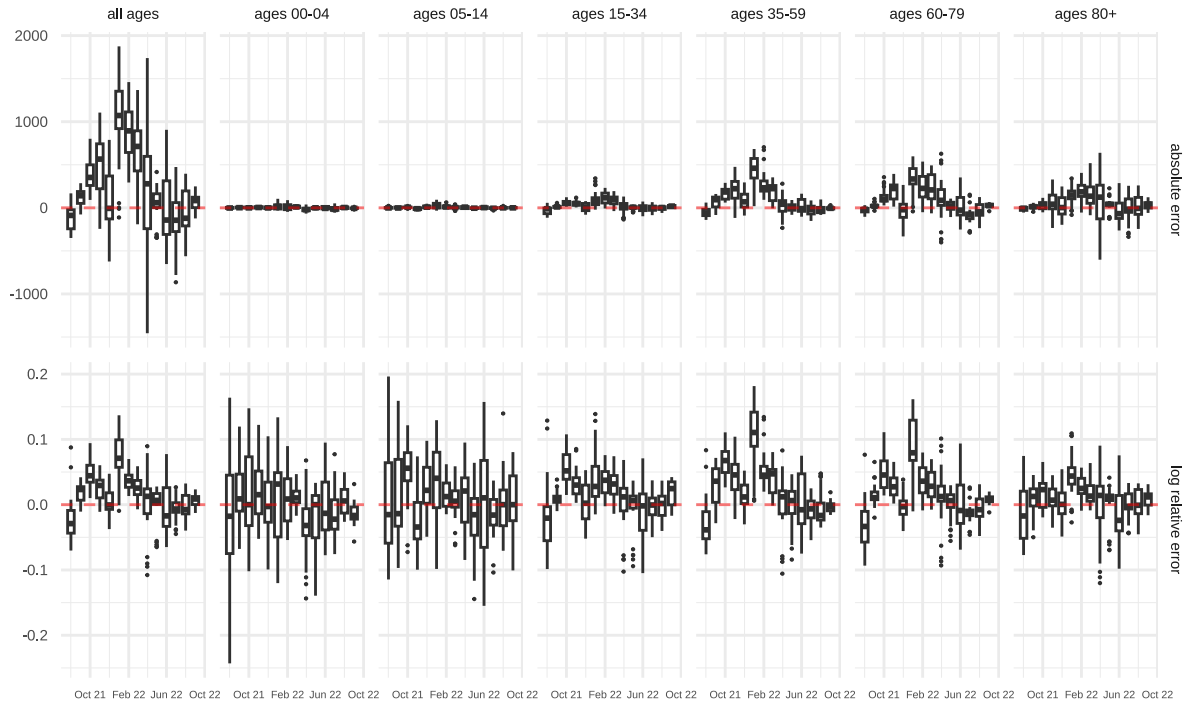

Figure 4: We show relative and absolute errors of prediction of our model for same day nowcasts by month of forecast and selected age groups. Relative errors are displayed on the log10 scale, i.e. as $\log_{10}(\text{predicted}) - \log_{10}(\text{actual})$. Up to December 2021 the model performs well, especially in the older age groups where most hospitalisations occur. The sharp increase in cases in January 2022 coupled with a lower probability of hospitalisation, most likely due to the appearance of the Omicron variant in Germany, lead to overpredictions across all age groups.

To investigate the quality of point predictions we display the time-evolution of absolute (AEP) and relative errors of predictions (REP, $\log_{10}$-scale) across all age groups in Figure 4. From this figure one can infer that the point nowcasts produced by our model tend to slightly overpredict the final number of hospitalisations. Indeed, the interquartile range of REPs for all age groups and dates combined spans $[-1.56, 8.33]$, demonstrating the same tendency. The highest REPs occured in October / November 2021 and January/February 2022; the first corresponding to introduction of mandatory testing at the workplace and the second to the arrival of the Omicron variant in Germany. In both circumstances the number of cases rose while hospitalisations did not increase proportionally, a similar effect to the one observed in Figure 3.

We further quantified uncertainties in estimaton by uncertainty intervals based on an assumption of a (log)-normal distribution for the errors with standard deviation based on past performance of our model. In Figure 5 we show the coverages of the 50% and 95% prediction interval across all age groups and delays for the whole time period of our study. For most age groups the 50% prediction interval has close to nominal coverage, while the 95% intervals have less than nominal coverage.

As our goal is to capture all of the uncertainty in this prediction, we chose to assume a sensible distribution for the prediction, a normal distribution for the two young age groups and a log-normal distribution for all other age groups. This has the advantage of producing more honest, wider, prediction intervals than those based on parametric distributions. The estimated standard deviation will also account for periods of low coverage, such as January 2022, albeit only after the maximum delay of 12 weeks.

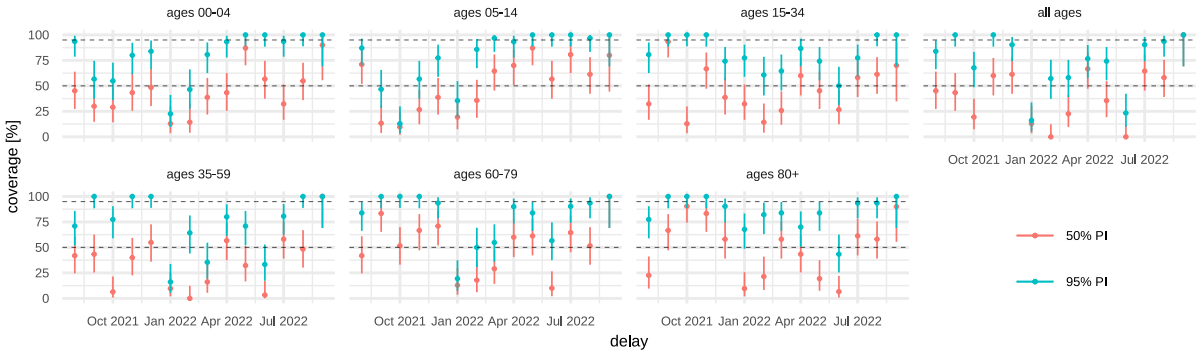

Figure 5: Empirical coverage of 50% and 95% prediction intervals (PI) based on same-day nowcasts for dates 2021-08-01 to 2022-09-10 (406 dates) for which the true amount of hospitalisations after 12 weeks is known as of the writing of this paper. We also display pointwise 95% binomial confidence intervals for the coverage. Given the difficulties of real-time forecasting [5] we deem the coverages good, except for the transitionary period in the end of 2021 where changing testing schemes and the change from Delta to Omicron cause our model to be overly confident. Coverage is generally better in the older age groups.

We base our choice of 12 weeks of delay on the empirical survival function displayed in Figure 7. One could, however, argue for shorter maximum delay such as 6 weeks because time from reported infection to hospitalisation is much shorter, on the order of ≈ 10 days [6], so hospitalisations after this (shorter) period are unlikely to be due to the acute infection with SARS-CoV-2. This would have two main benefits: The model would adapt faster to changing circumstances and the indicator nowcasted describes the severity of the epidemic more appropriately.

Discussion

The main advantage of our model over established nowcasting approches is its simplicity, making it easy to understand, straightforward to implement and, once the reporting triangles for incidences and hospitalisation are created, fast to run; taking only $< XX$ minutes on a standard notebook(**TODO: check!**).

The problem of nowcasting hospitalisations is different from previously studied nowcasting settings in several ways. At the time of nowcast a large fraction of hospitalisations are not only unobserved, but are yet to occur - in this sense the nowcast is more accurately termed a forecast. As the date of hospitalisation is not known, the hospitalisations are associated by the date of reporting of the COVID-19 case, creating the double-weekday effect displayed in 2. While daily updated data on hospitalisations are available, these consist only of moving weekly aggregates, consecutive observations are strongly auto-correlated.

We sidestep all of these issues by splitting the hospitalisations to nowcast into weekly chunks, incorporating leading indicators of hospitalisation – the weekly reported case incidences – and modelling the number of hospitalisations to come in each chunk by binomial thinning of incidences. Let us stress that this approach is only possible in the special situation where case and hospitalisation are explicitly linked, however we believe that incorporating leading indicators into nowcasting models is a promising approach.

An additional advantage of our model is that the hospitalisation probabilities can further be analysed, e.g. by investigating association between the publicly available vaccination rates and the probability of hosptialisation and delay to hospitalisation. Sudden changes in these fractions, as observed in Figure 3, can also hint towards worse model performance, especially if this change can be attributed to changing probability of hospitalisation due to new variants or changing testing regime.

Real time forecasting of epidemiological indicators is a difficult task [5], in particular quantifying uncertainty [4]. To test our model under real-time circumstances we submitted daily nowcasts to the German COVID-19 NowcastHub [16] since Novemer 2021. In the nowcasting context, [12] goes to great lengths to account for overdispersion due to changes in delay distribution, introducing gamma and Dirichlet priors and explicitly

modelling trends. Such an approach would also be feasible for our approach, e.g. model incidences by an appropriate Poisson or negative binomial distribution and, conditional on incidences, model hospitalisations by a binomial distribution. As this increases the complexity of our model and relies on the assumed distributions being sensible we opted for another approach.

Regarding the indicator we stress that its value on a given date does not represent the current occupancy of hospitals in Germany with COVID-19 patients but is rather an approximation to the morbidity caused by COVID-19 on that date. The reason for this discrepancy is that hospitalisations are attributed to the reporting date of the associated case, not that of hospitalisation. While the reporting date of the hospitalisation can be recovered from the publicly available data, the date of hospitalisation cannot. Additionally, no information on the duration of stay is available, making it impossible to create an indicator for the occupancy of hospitals based solely on data provided by the Robert Koch-Institut.

Implicit in all of these approaches is an assumption of "stationarity", i.e. that future reported hospitalisations will behave as they did in the past. Thus, all of these approaches might still be insufficient if circumstances change drastically, for example introduction of new testing schemes (school, 3G at workplace), changes in the delay distribution due to new variants, or hospitals close to capacity taking longer to process cases.

In summary, because models usually only capture a small part of the highly dynamic data-generating process, we believe that uncertainty in such circumstances should not come from unrealistic parametric assumptions but rather be based on past model performance. Given the discussed difficulties and the changing epidemiological dynamics in the period studied, the observed errors of prediction (Figure 4) and coverages of prediction intervals (Figure 5) are satisfying.

In this paper we provide a straight-forward model for nowcasting hospitalisations associated with COVID-19 in Germany. By leveraging known incident cases, we can estimate fractions of hospitalisations in weekly chunks which in turn avoids a complicated model of the two weekday effects present in the data. As the circumstances of the epidemic are changing constantly, e.g. vaccination coverage, testing regimes and emerging variants, we based uncertainty not on parametric assumptions but on the past performance of our model, assuming a (log)normal predictive distribution. We contributed nowcasts based on this model since November 2021 to the German COVID-19 NowcastHub [16], a collaborative platform collecting and aggregating such nowcasts from multiple research groups. The performance of the nowcasts in this Hub and presented in this paper (Figures 4 and 5) are, regarding the simplicity of the model and the highly dynamic situation, quite satisfying.

There are multiple extensions to our model worth investigating. Firstly hospitalisations are also available at the federal state level so nowcasting on a spatial scale is naturally of interest due to hetereogeneity in immunisation status and testing regimes across states. However, splitting hospitalisations into six age groups and 16 federal states will result in small numbers with larger variability which in turn increases variabliity in estimates $p_{t,d,k}$ and thus predictions which may require some regularisation thus increasing the models complexity. Secondly, in a similar vein, modeling the temporal evolution of hospitalisation probabilities by smooth functions, e.g. splines [24, 23], may help in early detection of changing circumstances and thus lead to better forecasts. Thirdly our uncertainty intervals account for the variance of past performance, but Figure 4 suggests that there is substantial bias in periods of changing circumstances which could be incorporated into our model in a straightforward way. Finally to predict the future course of the epidemic forecasting hospitalisations for dates t that lie in the future is of interest, which could be accomplished if one has a model that produces forecasts for incidences for all age groups.

Materials and Methods

To predict the number of hospitalisations we consider the reporting process of both reported COVID-19 cases and reported hospitalisations. Recall that the reporting date of a COVID-19 case is shared for both the case and its hospitalisation, i.e. the case and hospitalisation are linked through this date.

As hospitalisations are only available as 7-day rolling sums, see Section , we use 7-day rolling sums for daily reported incidences as well. To avoid dealing with the double weekday effect of both reporting date of the

case and reporting date of the hospitalisation (see 2) we divide the future hospitalisations we wish to predict into chunks of one week, which gets rid of the weekday effect for the hospitalisations. This is depicted in 6. Our prediction of each of these weekly chunks then consists of the fraction of hospitalisations of reported cases in the past.

Statistical Model

More formally, denote by $h_{t,d}$ the number hospitalisations with reporting date t that are known d days later. Unfortunately we only observe

$$H_{t,d} = \sum_{s=t-6}^{t} h_{s,d+(t-s)},$$

i.e. a weekly sum of reported hospitalisations. On day T our goal is to predict $H_{t,D}$ for large delays D and days $t \leq T$, of course it suffices to predict $H_{t,D} - H_{t,T-t}$ and add the known $H_{t,T-t}$ to this prediction. We rewrite this into weekly telescoping sum

$$H_{t,D} - H_{t,d} = (H_{t,d+7} - H_{t,d}) + (H_{t,d+14} - H_{t,d+7}) + \cdots + (H_{t,D} - H_{t,d+7K}),$$

where $K = \lfloor (D-d)/7 \rfloor$, reducing the task at hand to predict hospitalisations in the k-th week ahead, $H_{t,d+7k} - H_{t,d+7\cdot(k-1)}$, $k = 1, \ldots, K$. To leverage known reported incidences, rewrite this as

$$\underbrace{\frac{H_{t,d+7k} - H_{t,d+7\cdot(k-1)}}{I_{t,d}}}_{=:p_{t,d,k}} I_{t,d}$$

where $I_{t,d}$ is the 7-day case incidence with reporting date t known at time $t+d$, i.e. the incidenct case analouge of $H_{t,d}$.

Assuming that the proportions $p_{t,d,k}$ change slowly over time t we estimate them by

$$\widehat{p_{t,d,k}} = \frac{H_{t-7k,d+7k} - H_{t-7k,d+7\cdot(k-1)}}{I_{t-7k,d}} = p_{t-7k,d,k} \tag{1}$$

and finally predict

$$\widehat{H_{t,D}} = H_{t,d} + I_{t,d} \left(\widehat{p_{t,d,1}} + \cdots + \widehat{p_{t,d,K}} \right). \tag{2}$$

As hospitalisation is affected by age, we perform this procedure for all available age groups separately and finally aggregate over all age groups to obtain a nowcast for all age groups combined.

This describes our point nowcast for 7-day hospitalisations. To obtain uncertainty intervals we fit a normal (age groups 00-04 and 05-14) or lognormal (all other age groups) distribution to the past performance of our model. We chose these distributions based on explorative analysis and believe that these should be seen as heuristics rather than as a matter of fact, which is in line with the philosophy of our model to be as simple as possible.

Denote by $\hat{H}_{t,D,s}$ the nowcast made for date t on date $s \geq t$. Starting with date $t+D$ the definite $H_{t,D}$ is known and we can estimate the absolute prediction error $\varepsilon_{t,s} = H_{t,D} - \hat{H}_{t,D,s}$ and the relative prediction error $\eta_{t,s} = \log(H_{t,D} - H_{t,s-t}) - \log\left(\hat{H}_{t,D,s} - H_{t,s-t}\right)$. For the nowcast for date t made on date s we estimate the standard deviation $\hat{\sigma}$ of $\varepsilon_{t-D-i,s-D-i}$ or $\eta_{t-D-i,s-D-i}$ (age groups 00-04, 05-14 and others respectively), $i = 0, \ldots, 27$ by its empirical counterpart. The estimated predictive distribution which informs our prediction intervals is then $\mathcal{N}(\hat{H}_{t,D,s}, \sigma^2)$ (age groups 00-04 and 05-14) or $\mathcal{LN}\left(\log\left(\hat{H}_{t,D,s} - H_{t,s-t}\right), \sigma^2\right) + H_{t,s-t}$ (all other agr groups).

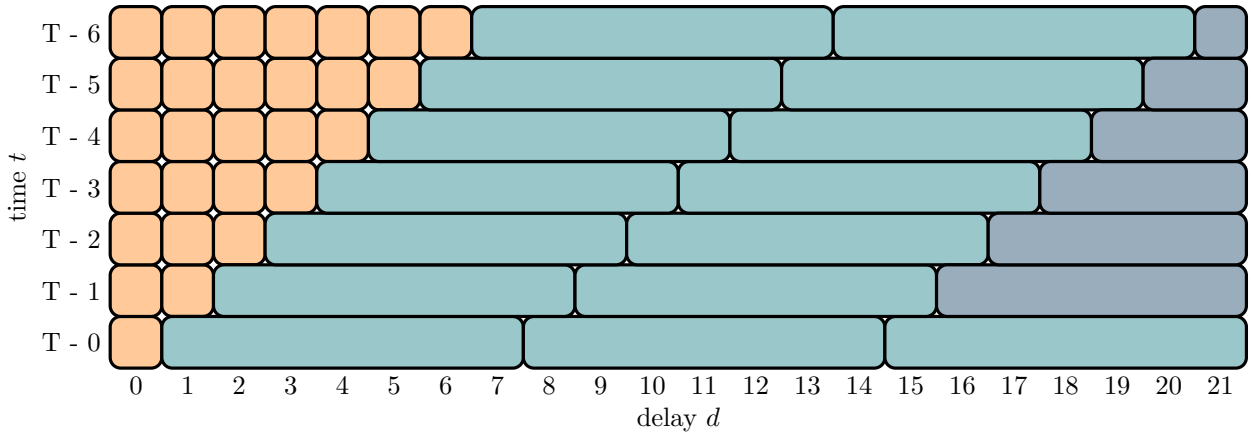

Figure 6: Decomposition of the daily reported hospitalisation incidences into the known incidences, i.e. the **reporting triangle**, and the future weekly increments. The last increment might not be a weekly one, but we expect few cases to occur for such long delays.

Data & Parameters

We use publicly available data from the German national health authority (Robert-Koch institut) on daily reported COVID-19 cases [20] and weekly reported hospitalisations [19]. Both datasets are updated on a daily basis.

COVID-19 cases are described by their date of reporting, i.e. the date that the local health authorities were made aware of the case. For a fraction (63 %) of cases the date of symptom onset is also reported. Due to delays in the process from infection to reporting – e.g. the time it takes to get tested, evaluate the test and report the result to local health authorities – the date of reporting is, for most cases, some days after symptom onset (median delay: 3 with interquartile range $[2, 6]$). As the date of symptom onset is not known for a substantial amount of incident cases, and is not reported for hospitalised cases, we focus our analysis on the date of reporting.

Hospitalisations are associated with the *reporting date of the corresponding case* and no information is available on the actual date of hospitalisation. In addition, hospitalisations are only published as weekly sums over the past seven days. This means that the number of hospitalisations reported for today consists of all hospitalisations that correspond to cases that have a *case reporting date* in the past seven days. In particular if the case reporting date of a hospitalised case is today the case will *not* count towards todays hospitalisation count. The reporting date of hospitalisation is not available in the dataset, but can be inferred by comparing datasets from consecutive days.

Daily incident cases and weekly hospitalisations are reported by federal state and age group (00-04, 05-14, 15-34, 35-59, 60-79, 80+). Incident cases are additionally reported by county and sex.

In line with the structure of the data provided by the Robert-Koch institut we let $H^a_{t,d}$ be the number of weekly hospitalisations in age group a with case reporting date $t-1, \dots, t-7$ that are known on day $t+d$, aggregating over all states. Accordingly we define $I^a_{t,d}$ to be the number of weekly incident cases in age group a with reporting date $t-1, \dots, t-7$ that are known on day $t+d$. Finally we reconstruct the reporting triangles for weekly hospitalisations (Figure 6) by differencing the $H^a_{t,d}$ for fixed t: $h^a_{t,d} = H^a_{t,d} - H^a_{t,d-1}$, setting $H^a_{t,-1}$ to 0 by convention. We recover the reporting triangle $i^a_{t,d}$ for incident cases in the same manner.

We show the empirical surivival function of hospitalisations for a fixed date in Figure 7. We observe that delays have long tails, with most cases reported after 12 weeks (84 days), except for the youngest age group. After such a long delay between infection and hospitalisation we deem it unlikely that hospitalisation is due to COVID and disregard all longer delays accordingly. Given such long delays, it does not suffice to nowcast only todays hospitalisations, but also for dates in the past to monitor hospitalisation, i.e. observe current trends; we thus nowcast for all delays $d = 0, \dots, 28$.

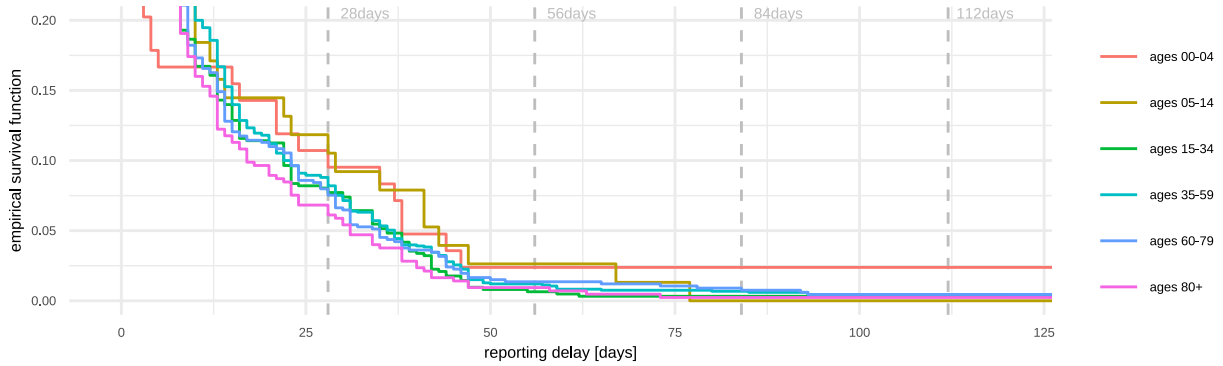

Figure 7: Survival function of reporting delays of weekly hospitalisations $H_{t,d}$ with case reporting date 01 September 2021. The delay distribution has long tails with a non-neglibgible fraction of observed delays longer than eight weeks, in some age groups even twelve weeks.

Additional points to make?

- despite difficulties as a monitoring tool for real-time-decision making the incidence is of epidemiological interest because it depicts how many hospitalisations were generated on that day (if you believe date of infection to be close to date of reporting) which might be less affected by changes in testing protocols
- also while date of hospitalisation is a late indicator, the date of the case is not, e.g. if we can correct for hospitalisation delay we have real time monitoring of severity of epidemic

Literature

[1] Andrei R. Akhmetzhanov. "Estimation of Delay-Adjusted All-Cause Excess Mortality in the USA: March-December 2020." In: *Epidemiology and Infection* (2021). DOI: 10.1017/s0950268821001527.

[2] Matthias An Der Heiden and Osamah Hamouda. "Schätzung Der Aktuellen Entwicklung Der SARS-CoV-2- Epidemie in Deutschland – Nowcasting". In: *Epidemiologisches Bulletin* (Apr. 2020). DOI: 10.25646/6692.4.

[3] Leonardo S Bastos et al. "A Modelling Approach for Correcting Reporting Delays in Disease Surveillance Data". In: *Statistics in Medicine* 38.22 (2019), pp. 4363–4377. ISSN: 1097-0258. DOI: 10.1002/sim.8303.

[4] J. Bracher et al. "A Pre-Registered Short-Term Forecasting Study of COVID-19 in Germany and Poland during the Second Wave". In: *Nature Communications* 12.1 (Aug. 2021), p. 5173. ISSN: 2041-1723. DOI: 10.1038/s41467-021-25207-0.

[5] Angel N. Desai et al. "Real-Time Epidemic Forecasting: Challenges and Opportunities". In: *Health Security* 17.4 (Aug. 2019), pp. 268–275. ISSN: 2326-5094. DOI: 10.1089/hs.2019.0022.

[6] Christel Faes et al. "Time between Symptom Onset, Hospitalisation and Recovery or Death: Statistical Analysis of Belgian COVID-19 Patients". In: *International Journal of Environmental Research and Public Health* 17.20 (Jan. 2020), p. 7560. ISSN: 1660-4601. DOI: 10.3390/ijerph17207560.

[7] C. P. Farrington et al. "A Statistical Algorithm for the Early Detection of Outbreaks of Infectious Disease". In: *Journal of The Royal Statistical Society Series A-statistics in Society* (1996). DOI: 10.2307/2983331.

[8] Felix Günther et al. "Nowcasting the COVID-19 Pandemic in Bavaria". In: *Biometrical Journal* 63.3 (2021), pp. 490–502. ISSN: 1521-4036. DOI: 10.1002/bimj.202000112.

[9] Michael Höhle and Matthias An Der Heiden. "Bayesian Nowcasting during the STEC O104:H4 Outbreak in Germany, 2011". In: *Biometrics* 70.4 (2014), pp. 993–1002. ISSN: 1541-0420. DOI: 10.1111/biom.12194.

[10] *Hospitalization Nowcast Hub*. KITmetricslab. Oct. 2022.

[11] Kenneth S. Kaminsky. "Prediction of IBNR Claim Counts by Modelling the Distribution of Report Lags". In: *Insurance Mathematics & Economics* 6.2 (Apr. 1987), pp. 151–159. DOI: 10.1016/0167-6687(87)90024-2.

[12] J. F. Lawless. "Adjustments for Reporting Delays and the Prediction of Occurred but Not Reported Events". In: *Canadian Journal of Statistics* 22.1 (1994), pp. 15–31. ISSN: 1708-945X. DOI: 10.2307/3315826.n1.

[13] Sarah F. McGough et al. "Nowcasting by Bayesian Smoothing: A Flexible, Generalizable Model for Real-Time Epidemic Tracking." In: *PLOS Computational Biology* (2020). DOI: 10.1371/journal.pcbi.1007735.

[14] Douglas Midthune et al. "Modeling Reporting Delays and Reporting Corrections in Cancer Registry Data". In: *Journal of the American Statistical Association* 100.469 (Mar. 2005), pp. 61–70. DOI: 10.1198/016214504000001899.

[15] Angela Noufaily et al. "Modelling Reporting Delays for Outbreak Detection in Infectious Disease Data". In: *Journal of The Royal Statistical Society Series A-statistics in Society* (2015). DOI: 10.1111/rssa.12055.

[16] *Nowcasts Der COVID-19 Hospitalisierungsinzidenz*. https://covid19nowcasthub.de/. 2022.

[17] Evan L. Ray et al. "Ensemble Forecasts of Coronavirus Disease 2019 (COVID-19) in the U.S." In: *medRxiv* (Aug. 2020), p. 2020.08.19.20177493. DOI: 10.1101/2020.08.19.20177493.

[18] Arthur E. Renshaw and Richard J. Verrall. "A Stochastic Model Underlying the Chain-Ladder Technique". In: *British Actuarial Journal* 4.4 (1998), pp. 903–923. DOI: 10.1017/S1357321700000222.

[19] Robert Koch-Institut. *COVID-19-Hospitalisierungen in Deutschland*. Feb. 2022. DOI: 10.5281/ZENODO.5519056.

[20] Robert Koch-Institut. *SARS-CoV-2 Infektionen in Deutschland*. Feb. 2022. DOI: 10.5281/ZENODO.4681153.

[21] Robert Koch-Institut. *Wöchentlicher Lagebericht Des RKI Zur Coronavirus-Krankheit-2019 (COVID-19)*. Jan. 2022.

[22] Maëlle Salmon et al. "Bayesian Outbreak Detection in the Presence of Reporting Delays". In: *Biometrical Journal* (2015). DOI: 10.1002/bimj.201400159.

[23] Marc Schneble et al. "Nowcasting Fatal COVID-19 Infections on a Regional Level in Germany". In: *Biometrical Journal* 63.3 (Mar. 2021), pp. 471–489. ISSN: 0323-3847, 1521-4036. DOI: 10.1002/bimj.202000143.

[24] Jan van de Kassteele et al. "Nowcasting the Number of New Symptomatic Cases during Infectious Disease Outbreaks Using Constrained P-Spline Smoothing". In: *Epidemiology (Cambridge, Mass.)* (2019). DOI: 10.1097/ede.0000000000001050.

[25] Scott L. Zeger, Lai-Chu See, and Peter J. Diggle. "Statistical Methods for Monitoring the AIDS Epidemic". In: *Statistics in Medicine* 8.1 (1989), pp. 3–21. DOI: 10.1002/sim.4780080104.