SUMMARY SHORT-TERM PREDICTIONS

1. Model INLA (Toon Braeye, Sciensano)

Different methods exist to predict trends in epidemics. The method used by Sciensano focuses on observed trends and delays between outcomes of different indicators to predict the evolution of the number of new cases and new hospital admissions on a short-term period.

First, the time-varying reproduction number (R_t) is estimated by (1) the number of persons reporting symptoms (from test prescription data) (2) the number of incident cases in the nursing homes, (3) the number of cases overall (lab-confirmed PCR and antigen tests), (4) the hospital admissions for COVID-19 (referrals between hospitals are not included). Furthermore, information on the total number of tests performed (lab-confirmed, negative and positive PCR and antigen tests) is included.

The relation between these different R_t's is then established by distributed-lag non-linear regression models. The model with the best fit (lowest AIC[1]) based on fitting the model to the data that were collected during the last two months, is then selected to use for the prediction of the hospital intakes. Values that are not yet observed but are necessary for prediction, are assumed to have a stable trend over the last week of observed data and the unobserved period.

To predict future hospital admissions, the most recent observed numbers are combined with the predicted R_t-values. The above described process is done both for the data collected by province and by age group.

This method has different limitations. Since predictions are based on the observed data, recent measures or events that have not yet resulted in a trend change will not affect the predictions.
When more observed data are used instead of data based on the assumption of a stable trend, we expect the prediction to be more accurate. The relevant delay is estimated by the model.

2. Model GAM (U Hasselt (C.Faes, I-BioStat, Hasselt University)

The short-term prediction model of UHasselt is based on a statistical regression model, called a distributed lag non-linear model ([2]). The model compares the trend in the number of hospital admissions with a set of early-warning predictors of hospital admission.
Early warning predictors are indicators of which the observed value on a given day, is related to the number of new hospitalizations some days later. These indicators that serve as predictors are pre-smoothed to take out day-to-day variability using spline smoothing models ([3] [4]). The best predictors for new hospitalizations are the test positivity ratio of the past week and the mobility. The mobility is based on mobile network data, and is highly correlated with the intervention measures taken. To allow for a prediction for a period of 2 weeks, the model is then further informed with the number of patients with respiratory symptoms that visit the general practitioner (barometer 2.0 data) and absenteeism at work, which is associated with new hospitalizations in 10 to 14 days. The model provides the best predictions in the short-term (1 to 2 weeks). Uncertainty in the prediction inevitably increases for predictions further in time. The model is calibrated daily on provincial data, based on all available information in the number of hospitalizations and in the predictors since March 15.

[1] Akaike information criterion
[2] Gasparrini, A., Scheipl, F., Armstrong, B., Kenward, M. G. (2017). A penalized framework for distributed lag non-linear models. Biometrics, 73(3), 938-948.
[3] Aerts, M., Hens, N., Simonoff, J. S. (2010). Model selection in regression based on pre-smoothing. Journal of Applied Statistics, 37, 1455-1472.
[4] Wood S.N. (2017) Generalized Additive Models: An Introduction with R (2nd edition). Chapman and Hall/CRC Press.