

Toelichting schatting reproductiegetal

Han-Kwang Nienhuys (2021-08-01; Twitter: @hk_nien)

Het verband tussen groei en krimp van de dagelijks gerapporteerde Covid-19-infecties en het reproductiegetal R wordt bepaald door:

- het generatie-interval $T_c = 4$ d: de tijd tussen het krijgen en doorgeven van een infectie. Dit is de waarde waar RIVM mee rekent. (RIVM leek tot 1 juli 2021 exact 4 dagen aan te houden, zonder spreiding.)
- De gemiddelde tijd T_{ir} tussen het moment van infectie en het verschijnen in de dagcijfers, momenteel ca. 6 dagen.

Als $I(t)$ het aantal infecties op dag t is, dan $I(t) = R(t) I(t - T_c)$, waarin $R(t)$ het reproductiegetal op dag t is. Als R constant is, dan zie je exponentiële groei in $P(t)$, of een lineaire groei van de $\ln P$:

$$\frac{d \ln P}{dt} = \frac{\ln R}{T_c}.$$

Als R varieert in de tijd, dan

$$\ln R \left(t + \frac{T_c}{2} \right) = \ln \left(\frac{I(T + T_c/2)}{I(T - T_c/2)} \right) \approx T_c \frac{d}{dt} \ln I(t) = T_c \frac{d}{dt} \ln P(t + T_{ir}).$$

De rapportagevertraging, d.w.z. de vertraging tussen veranderingen in R en veranderingen in de groei van P is dus $T_r = T_{ir} - T_c/2$, dus 4 dagen.

Hierop zijn twee correcties nodig:

- De dagcijfers P schommelen nogal met de dag van de week. Dit is te verhelpen door het gemiddelde P_7 te gebruiken van 3 dagen eerder t/m 3 dagen later (7-daags gemiddelde).
- De rapportagevertraging is niet helemaal constant: momenteel 4 dagen, maar rond 1 oktober 9 dagen. Als $T_r(t)$ de vertraging was voor rapportage op tijdstip t , dan is het reproductiegetal

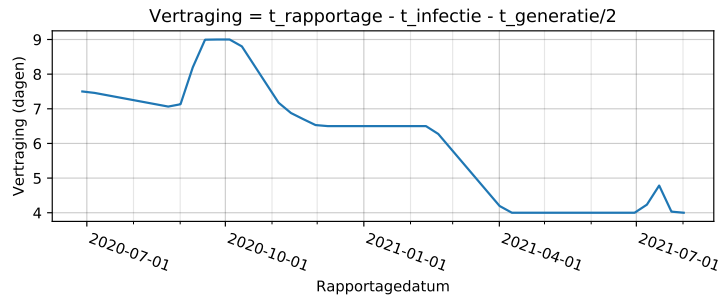
$$R(t - T_r(t)) = \exp \left(T_c \frac{d}{dt} \ln \left(\frac{P(t)}{1 - dPr/dt} \right) \right).$$

Hiermee kun je het reproductiegetal schatten tot 8 dagen terug (4 voor de rapportagevertraging, 3 voor het 7-daags-gemiddelde, en 1 voor de afgeleide d/dt). Deze methode levert binnen een marge ± 0.06 dezelfde waarden op als RIVM, maar

RIVM publiceert ze slechts 2x per week tot 14 dagen terug (voorheen 1x per week met 18 dagen vertraging).

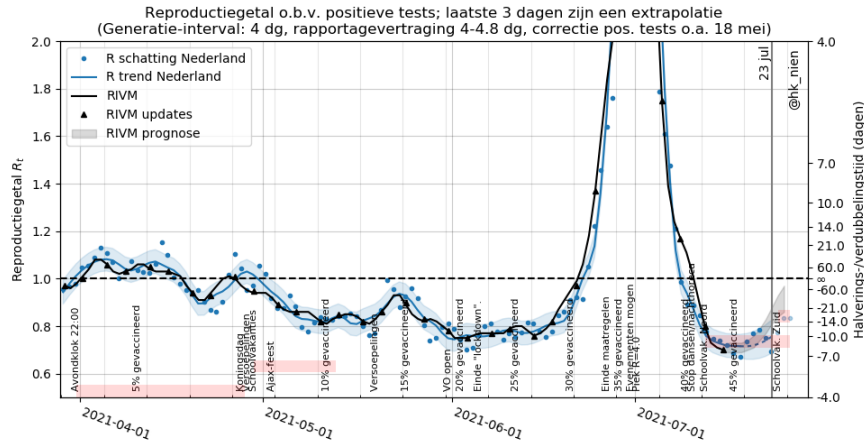
De open-source code staat op <https://github.com/han-kwang/covid19>.

Dit is de rapportagevertraging $T_r(t)$ die ik eind juli 2021 gebruik:



De tijdelijke toename rond 15 juli 2021 is een schatting vanwege berichten over capaciteitsproblemen in enkele laboratoria.

Een voorbeeld van een reproductiegetal-grafiek:



De zwarte curve is van RIVM (zwarte stippen voor de wekelijkse update). De grijze band is de prognose van RIVM (die gewoonlijk weinig lijkt op de definitieve cijfers een week later). De RIVM-curve is systematisch lager op dinsdagen en hoger op vrijdagen/zaterdag.

Rond 1 juli 2021 (grote besmettingsgolf) lijkt RIVM de rekenmethode te hebben gewijzigd (generatieinterval 3.7 dagen of grote spreiding in generatieinterval in plaats van exact 4.0 dagen), waardoor RIVM $R = 3.0$ vond en ik met bovenstaande rekenmethode $R = 4.0$ vond.

De blauwe stippen zijn de berekening zoals hier toegelicht. Op de meest recente drie punten na zijn die definitief. (Tenzij er wegens IT-problemen data niet is)

doorgekomen, zoals rond 16 december). De blauwe lijn is een vloeiende curve; de gestippelde laatste zes dagen van die curve zijn niet definitief.