DE CORONA CODE

De oorlog tegen het coronavirus woedt op vele fronten. Hij wordt gevoerd door patiënten, door medisch personeel, en door de gehele bevolking die *social distancing* in acht neemt. Ook statistici en epidemiologen voeren oorlog. Aan de hand van computermodellen proberen zij te voorspellen hoe de ziekte zich verder zal ontwikkelen: hoe besmettelijk zij is, hoeveel patiënten nog op de intensive care terecht zullen komen, en wat het effect is van specifieke beleidsmaatregelen.

Een recent nieuwsitem in *Nature*[1] doet verslag van een team epidemiologen van Imperial College London, geleid door Neil Ferguson. Hun computermodel voorspelde dat, als het Verenigd Koninkrijk niet handelend zou optreden, er mogelijk een half miljoen Britten zou overlijden. Naar verluidt zette deze voorspelling Boris Johnson ertoe aan drastisch van strategie te veranderen.

Modellen zoals die van Ferguson spelen dus een grote rol bij het bepalen van beleid rondom de coronacrisis. Gegeven het publieke belang van deze modellen mag je verwachten dat de onderliggende computercode publiekelijk beschikbaar is. Niet dus! In een tweet op 22 maart (zie rechtsonder) legde Ferguson uit waarom niet. Vanuit persoonlijk perspectief van Ferguson wellicht begrijpelijk, vanuit maatschappelijk perspectief is het ronduit beangstigend. Er zal maar een belangrijke fout in die dertien jaar oude niet-gedocumenteerde computercode zitten …

Dat zo'n fout niet onmogelijk is, blijkt bijvoorbeeld uit de beruchte affaire-Reinhart-en-Rogoff: twee economen die in 2010 dachten aangetoond te hebben dat landen met een hoge staatsschuld een lagere economische groei laten zien. Destijds spoorde dit inzicht politici aan om, in een tijd van economische crisis, toch verregaande bezuinigingen door te voeren. Maar wat bleek? Reinhart en Rogoff hadden een Excel-blunder begaan – ze hadden verkeerd met de muis gesleept, waardoor sommige cellen niet geselecteerd waren. Deze blunder bleek cruciaal: met correct uitgevoerde berekeningen waren de resultaten niet eenduidig, en boden ze helemaal geen steun voor een streng bezuinigingsbeleid tijdens crisistijd.

Het is dus niet onmogelijk dat de code van Ferguson grote fouten bevat, maar het is gelukkig wel onwaarschijnlijk: andere teams van epidemiologen kwamen tot vergelijkbare voorspellingen. Maar het feit dat de onderliggende code niet beschikbaar is, betekent wel dat het Ferguson-team zich afzondert van de rest van de wetenschappelijke wereld – die dus ook niet kan bijdragen aan verdere ontwikkeling en verfijning van het model, laat staan dat andere landen het model kunnen gebruiken. Gelukkig liet Ferguson weten dat het publiek maken van goed gedocumenteerde code een van zijn hoogste prioriteiten is.

En ons eigen Regionaal Instituut voor Veiligheid en Milieu, hoe staat het met de code voor hun computermodellen? Vooralsnog lijkt die code goed opgesloten in de RIVM-kluis, veilig voor de buitenwacht. Laten we hopen dat het niet gaat om duizenden regels niet-gedocumenteerde code, maar zeker weten zullen we het allicht nooit.[2]

1. Adam, D. (2020). Special report: The simulations driving the world's response to COVID-19: How epidemiologists rushed to model the coronavirus pandemic. *Nature*. https://www.nature.com/articles/d41586-020-01003-6
2. Deze column is geïnspireerd door een discussie op het Slack-kanaal van https://dataversuscorona.com/.



Eric-Jan Wagenmakers

is hoogleraar Mathemathische Psychologie aan de Universiteit van Amsterdam.
E-mail: ej.wagenmakers@gmail.com

FOTO: PETER VALCKX

