



## Effectiveness of a COVID-19 contact tracing app in a simulation model with indirect and informal contact tracing

Ka Yin Leung<sup>a,\*</sup>, Esther Metting<sup>b,c,d</sup>, Wolfgang Ebbers<sup>e</sup>, Irene Veldhuijzen<sup>a</sup>, Stijn P. Andeweg<sup>a</sup>, Guus Luijben<sup>f</sup>, Marijn de Bruin<sup>f,g</sup>, Jacco Wallinga<sup>a,h</sup>, Don Klinkenberg<sup>a</sup>

<sup>a</sup> National Institute for Public Health and the Environment, Centre for Infectious Disease Control, Bilthoven, the Netherlands

<sup>b</sup> University of Groningen, University Medical Center Groningen, Data Science Center in Health, the Netherlands

<sup>c</sup> University of Groningen, University Medical Center Groningen, Department of Primary Care, the Netherlands

<sup>d</sup> University of Groningen, faculty of Economics and Business, Department of Operations, the Netherlands

<sup>e</sup> Erasmus School of Social and Behavioural Sciences, Department of Public Administration and Sociology, Erasmus University Rotterdam, Rotterdam, the Netherlands

<sup>f</sup> National Institute for Public Health and the Environment, Centre for Health and Society, Bilthoven, the Netherlands

<sup>g</sup> Radboud University Medical Centre, Radboud Institute of Health Sciences, IQ Healthcare, Nijmegen, the Netherlands

<sup>h</sup> Leiden University Medical Centre, Department of Biomedical Data Sciences, Leiden, the Netherlands

### ARTICLE INFO

#### Keywords:

Non-pharmaceutical interventions  
Digital contact tracing  
Branching process  
Clustering coefficient  
Behaviour  
Transmission model

### ABSTRACT

During the COVID-19 pandemic, contact tracing was used to identify individuals who had been in contact with a confirmed case so that these contacted individuals could be tested and quarantined to prevent further spread of the SARS-CoV-2 virus. Many countries developed mobile apps to find these contacted individuals faster. We evaluate the epidemiological effectiveness of the Dutch app CoronaMelder, where we measure effectiveness as the reduction of the reproduction number  $R$ . To this end, we use a simulation model of SARS-CoV-2 spread and contact tracing, informed by data collected during the study period (December 2020 - March 2021) in the Netherlands. We show that the tracing app caused a clear but small reduction of the reproduction number, and the magnitude of the effect was found to be robust in sensitivity analyses. The app could have been more effective if more people had used it, and if notification of contacts could have been done directly by the user and thus reducing the time intervals between symptom onset and reporting of contacts. The model has two innovative aspects: i) it accounts for the clustered nature of social networks and ii) cases can alert their contacts informally without involvement of health authorities or the tracing app.

### 1. Introduction

In most European countries, the main non-pharmaceutical control measures against the spread of SARS-CoV-2 in the first two years were aimed at reducing contacts or risk of transmission per contact on population level, for instance by closing schools or restaurants, working from home, keeping distance, or wearing face masks. Some of these measures may have had a negative impact on the society, in particular to mental health and the economy (Ashraf, 2020; Van den Boom, 2022). For that reason, control measures focusing on isolating only infected individuals and quarantining their contacts were preferable.

In infectious disease control, contact tracing is a widely applied strategy to find potentially infectious individuals, quarantine and test them. Traditionally, contact tracing is done manually by public health

professionals and relies on the ability and willingness of an individual to recall and report close contact events (to be referred to as contacts), and of the contacted persons (to be referred to as contactees) to follow the quarantine and testing rules. Modelling studies show that the effectiveness of contact tracing strongly depends on characteristics of the infection such as the incubation period and generation time, and can be very sensitive to delays in the tracing process such as waiting for test appointments or the time it takes for health authorities to call cases, ask for contacts, and then inform contactees, which increase the time between transmission and starting quarantine (Fraser et al., 2004; Hellewell et al., 2020; Klinkenberg et al., 2006; Kretzschmar et al., 2021; Muller et al., 2000).

To improve the process of contact tracing, various countries developed contact tracing apps which informed users that they were

\* Corresponding author.

E-mail address: [ka.yin.leung@rivm.nl](mailto:ka.yin.leung@rivm.nl) (K.Y. Leung).

<https://doi.org/10.1016/j.epidem.2023.100735>

Received 15 May 2023; Received in revised form 17 November 2023; Accepted 8 December 2023

Available online 12 December 2023

1755-4365/© 2023 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

contacted by someone who was tested positive for SARS-CoV-2. These apps were meant to reduce the time interval between symptom onset of a case and the identification of their contacts, and to trace contactees not known to or recalled by the case. Early modelling studies (Cencetti et al., 2021; Ferrari et al., 2021; Kucharski et al., 2020; Kurita et al., 2021) suggested a potentially relevant contribution of such apps to COVID-19 control. Post-implementation analyses in the UK showed that many infections were indeed prevented by their contact tracing app (Ferretti et al., 2020). However, also with the digital contact tracing app, control by social distancing measures remained necessary.

Since the start of the pandemic many studies have been conducted on contact tracing with contact tracing apps (Braithwaite et al., 2020; Jenniskens et al., 2021). Mathematical modelling studies have applied several types of models, ranging from compartmental models (Ge et al., 2021; Nuzzo et al., 2020) and branching process models (Bradshaw et al., 2021; Kretzschmar et al., 2020), to agent-based models (Cencetti et al., 2021; Kucharski et al., 2020; Quilty et al., 2021; Scott et al., 2021). A consistent conclusion across studies is that contact tracing apps can contribute to the control of an epidemic, but the extent of the impact is very sensitive to the percentage of the population using the app and the delays in the process of tracing and isolating infectious contacts. Some studies (e.g. (Kretzschmar et al., 2020)) emphasize that short delays of steps in the contact tracing process are important for effective tracing, whereas other studies (e.g. (Kucharski et al., 2020)) stress that effective tracing requires a high proportion of cases to isolate, and of their contacts to be traced.

In 2021, the National Institute for Public Health and the Environment (RIVM) was asked for a model-based evaluation of the epidemiological effectiveness of the Dutch contact tracing app CoronaMelder. Here we present the model and the evaluation. We measure effectiveness of contact tracing as the resulting reduction in the reproduction number  $R$ , which is defined as the mean number of secondary cases per primary case (see (Klinkenberg, 2021) for the official report in Dutch). The model is informed with data from surveillance sources and behavioural surveys during the first months of 2021 in the Netherlands. The model itself includes two features that are novel. First, we account for the clustered nature of social contact networks, which is relevant as it increases the probability that infected individuals can be notified by a case in their social network without having been infected by that case. Second, we explicitly account for notification of contacts by cases themselves, without involvement of health authorities or a tracing app. We estimate the additional effect of manual contact tracing over this informal notification, and the additional effect of the contact tracing app over the combined effect of informal notification and manual contact tracing.

## 2. Methods

### 2.1. Data sources

Model parameters were estimated from data on the Dutch COVID-19 epidemic as much as possible. Epidemiological data from three surveillance systems were used:

- Osiris: the national case notification registry containing all confirmed COVID-19 cases (Ward et al., 2005)
- CoronIT: tests carried out by the Municipal Health Services, from all public testing facilities
- HPZone: all contact tracing data

Behavioural data from three large studies were used:

- CGU: questionnaires about beliefs and adherence to measures, ran by the Corona Behavioural Unit of the RIVM (RIVM, 2021)
- LISS: questionnaires about the tracing app, beliefs and adherence to measures, with an existing longitudinal Social Sciences cohort panel (Van der Laan, 2021)

- PanelClix: questionnaire about the tracing app and how it is used, with an existing internet panel (Ebbbers, 2021)

### 2.2. Epidemiological situation

The evaluation was done for the period between 1 December 2020 and 31 March 2021, when vaccination was not yet available, and lockdown restrictions were in place such as keeping distance, working from home, and closed businesses. During that period, all testing was done by public testing facilities (no self-administered antigen tests). People were allowed to get tested when they had symptoms, or five days after contact with a positive case identified by manual contact tracing. Notification of close contacts through the contact tracing app was done through the health authorities: when cases were approached by the health authorities for manual contact tracing, they were asked whether they used the app, and if so, they were provided a key by which they could activate the app to notify their contactees. Notified individuals were advised by the app to quarantine.

### 2.3. Model formulation and analysis

The effectiveness of contact tracing was expressed in terms of a percentual reduction of the reproduction number  $R$  with contact tracing compared to the reference reproduction number without testing and tracing. The effectiveness is between 0 % and 100 % where 0 % corresponds to no effect of contact tracing and 100 % to full control of the epidemic. The reproduction number was determined with a simulation model for SARS-CoV-2 transmission and contact tracing. This was done for several parameter sets: the baseline parameter set, parameter sets for sensitivity analyses, and parameter sets for scenarios to explore improvements to the app and its use.

For each parameter set we simulated 10,000 transmission trees (outbreaks starting with one index) of at most 10 generations which is equivalent to about 1.5 months of virus spread starting with an incidence rate of 2500 cases during four days (with a mean generation interval of 4 days). In the simulations all cases and transmission links are known, therefore it is possible to calculate the mean number of secondary cases across all transmission chains, which is the reproduction number  $R$ . The epidemiological effectiveness of testing, case isolation, contact tracing and quarantine was determined by applying sets of control measures to the 10,000 simulated transmission trees, removing cases if they would have been prevented, and recalculating the reproduction number.

The control measures were applied at different levels:

1. no control
2. testing and informal tracing: cases test and go into quarantine because of symptoms; cases go into isolation because of a positive test result; contactees test and go into quarantine because of informal contact tracing
3. manual tracing: contactees test and go into quarantine because of manual contact tracing (in addition to 1)
4. tracing app: contactees test and go into quarantine because of an app notification (in addition to 1 and 2)

The total effectiveness for a parameter set was determined as the reduction in  $R$  from no control to control level 3. The contributions of informal tracing, manual tracing, and app-based tracing separately were determined as the reductions in  $R$  from no control to level 1, from level 1 to level 2, and from level 2 to level 3, respectively.

Specific details of the Dutch testing policy were modelled explicitly: in control level 1, testing of contacts was only done when they were symptomatic and in control levels 2 and 3, testing of contacts was done when symptomatic or five days after infection (whichever was earlier). Of course, in all cases testing, isolation and quarantine were also conditional on adherence. See also Section 2.3.2.

The simulation model consisted of three layers, explained in more detail below. The first layer was the transmission model, used to create transmission trees without contact tracing interventions. This resulted in a chronological description for all cases of who had contacted whom, and when symptoms had started (if any). The second layer was the behaviour model, used to assign behavioural parameters to each case that determine what actions they are willing to take to help control the epidemic. Actions are testing or isolation/quarantine, and are only taken when triggered by symptom onset, a positive test result, or notification through contact tracing. The behaviour model is also used to determine who is willing to use the contact tracing app (app use). The third layer was the logistics model, used to apply testing and tracing to the transmission trees with all relevant delays in the test-and-trace process, and the probabilities by which contacts are traced. The second- and third layer were tailored to the Dutch situation in the study period, e.g. testing rules and notification processes. Further details are provided in the model description below.

2.3.1. Transmission: epidemic outbreak and social network with clustering

Transmission trees were simulated with a branching process with negative binomial offspring distribution, up to a maximum number of 10 generations after the index case. The offspring distribution is defined by a mean  $R_{ref}$  (the reference reproduction number, i.e. without testing and tracing) and a shape  $k$ , which determines the variation and therefore the possibility of superspreading. The reproduction number  $R_{ref}$  in the simulation is without isolation or quarantine, but should be interpreted in the presence of population-level control measures such as keeping physical distance, working from home, and closed businesses.

For each case a time of infection was sampled, as well as a time of symptom onset, with a certain fraction of cases without symptoms (see also Table 1). Starting with infection time 0 for the index case of the branching process, the infection times for the subsequent cases were always one random generation interval later. The times of symptom onset were equal to the infection times plus a randomly sampled incubation period. Generation interval and incubation period were modelled with gamma distributions.

We accounted for the local clustering of social contact networks by adding edges to the transmission trees that represent contacts that had not led to transmission but which may be used for contact tracing. These edges were placed such that networks with triangles were formed, in three steps:

- each pair of individuals that had a contactee in common was linked with probability  $c$ .
- each pair of individuals that had a contactee in common because of step 1, was linked with probability  $c^2$ .
- Each pair of individuals that had a contactee in common because of step 2, was linked with probability  $c^3$ .

2.3.2. Behaviour: adherence to control measures

Behaviour was parameterised by a set of probabilities of adherence to testing, isolation, quarantine and app use (the behaviours), in response to symptoms, a positive test result, any type of contact tracing, or availability of a contact tracing app (the triggers). All individuals in the simulated outbreaks were randomly assigned three adherence levels between 0 and 1: one level related to testing, one related to isolation and quarantine, and one related to app use. These levels determined the behaviour of an individual in response to a trigger, by comparing their adherence level to a probability. For instance, if app use was 30 %, the 30 % individuals that were assigned the highest app-use adherence level (app-use level > 0.7) used the contact tracing app. The parameter app-use thus determines the percentage of the population that have the contact tracing app, which impacts both the percentage of the population that can notify their contacts as well as the number of contacts that can be reached through the app as both the case and his/her contacts need to use the app. By using adherence levels instead of random behaviours for each trigger-behaviour combination, we created consistency in each individual’s behaviour across triggers and parameter sets. The probabilities for each trigger-behaviour combination are listed in Table 2, with more detail in Supplement S1.

2.3.3. Logistics: delays and tracing efficiency

The process of testing and informing contacts, potentially followed by quarantine and isolation, is described by a cycle of two steps, initiated by a trigger step. Most actions are described conditional on adherence, as described above.

1. The trigger step is the period between symptom onset and scheduling a test. During this step, there is no quarantine.
2. The first step of the cycle is the period between scheduling the test and receiving the test result. During this period, the actual test is taken (not explicitly modelled). This first step may be the start of the

Table 1  
Parameters of the transmission model.

Description	Baseline value <sup>a</sup>	Source <sup>b</sup>
Reproduction number $R_{ref}$	1.3	<sup>c</sup>
Contacts		
Superspreading coefficient $k$ (shape parameter of the negative binomial distribution with mean $R_{ref}$ )	0.1	(Kirkegaard and Sneppen, 2021; Susswein, 2020)
Clustering probability $c$	0.2	(Eames and Keeling, 2003; Kissler, 2020; Kucharski et al., 2018)
Course of an individual infection		
Percentage of infections with symptoms	70 %	(Buitrago-García et al., 2020; McDonald et al., 2021)
Mean incubation period	5 days	(Backer et al., 2020; Cheng, 2020)
Mean generation interval	4 days	Osiris

<sup>a</sup> Value of the baseline parameter set; see Supplement S2 for values used in sensitivity analyses and scenarios

<sup>b</sup> Literature reference or dataset: Osiris = case notification registry; see Supplement for more detail

<sup>c</sup> The exact value of  $R$  is not essential to calculate a relative reduction. A value larger than one is necessary for an outbreak to enable computations. A value close to one is most representative for the evaluation period.

Table 2  
Parameters of the behaviour model. Adherence levels of actions in response to triggers.

Description	Baseline value <sup>a</sup>	Source <sup>b</sup>
Scheduling a test in response to...		
... symptoms (own initiative)	50 %	CGU
... informal tracing notification	68 %	CGU
... manual tracing notification	90 %	CGU
... tracing app notification	81 %	CGU, LISS, PanelClix
Isolation in response to a positive test	90 %	CGU
Quarantine in anticipation of test outcome, in response to...		
... symptoms (own initiative)	50 %	CGU
... informal tracing, manual tracing or tracing app (with symptoms)	75 %	CGU
... informal tracing, manual tracing or tracing app (without symptoms)	50 %	CGU, PanelClix

Parameter values with justification are listed in Table 1, with more detail in Supplement S1.

<sup>a</sup> Value of the baseline parameter set; see Supplement S2 for values used in sensitivity analyses and scenarios

<sup>b</sup> Dataset: CGU = questionnaires by Corona Behavioural Unit; LISS = questionnaires in LISS Social Sciences cohort panel; PanelClix = questionnaire in PanelClix internet panel; see Supplement for more detail

case's quarantine. Upon receiving the positive test result, the case may become an index in the tracing process.

3. The second step of the cycle is the period between receiving the positive test result of the index and the scheduling of tests by his/her contacts. This second step may be the start of isolation of the index. During this period, the contacts may be informed by the modes of contact tracing that are in place, with probabilities depending on the effectiveness of recalling and notifying by the index, the contact tracing app, and the authorities.

The parameter values of the logistics model are listed in Table 3, with more detail in Supplement S1.

#### 2.3.4. Applying testing and tracing to the simulated outbreaks

Implementation of control was simulated by first starting the test-and-trace cycle once for each case  $i$ , resulting in quarantine and isolation times  $t_{ij}$  for all cases  $j$  ( $t_{ij} = \infty$  if  $j$  is never reached from  $i$ ). From these times,  $\min_i(t_{ij})$  was determined as the end of the contact period for case  $j$  (i.e. minimum over all cases  $i$ ), and all secondary cases of case  $j$  after these end times were removed, plus all cases from the generations thereafter. From all resulting transmission trees, the numbers of cases  $C_g$  in each generation  $g$  were added, and the geometric average reproduction ratio was calculated over 7 generations as  $R = (C_9/C_2)^{1/7}$ ; we skipped the initial and final generation to reduce edge effects.

In sensitivity analyses and scenarios in which only parameters concerning behaviour or logistics were changed (Tables 2–3), the same set of 10,000 transmission trees and individual adherence levels were used as with the baseline parameter set. Only when parameters for the transmission tree model were changed (Table 1), new transmission trees were simulated.

#### 2.3.5. Baseline parameter set, sensitivity analyses, and scenarios

The baseline parameter set (Tables 1–3) was estimated from data or obtained from literature, and was meant to best represent the situation in the Netherlands between December 2020 and March 2021, as

**Table 3**  
Parameters of the logistics model.

Description	Value <sup>a</sup>	Source <sup>b</sup>
Delay from symptom onset to scheduling a test (own initiative)	1.6 days	Osiris, CoronIT
Delay from scheduling a test to test result... ...with symptoms	1.3 days	CoronIT
...without symptoms	2.3 days	CoronIT
Delay from positive test result to scheduling a test of a contactee, found by...		
...informal tracing notification	4 h	
...manual tracing notification	1.5 days	HPZone
...tracing app notification	1 day	
Percentage of infected contactees notified through...		
...informal tracing	32 %	
...manual tracing	40 %	Osiris
...tracing app	75 %	(Ministry of Health, 2020, 2021a)
Percentage of app users	16 %	PanelClix, (Ministry of Health, 2021b)

The code implementation of the model described in this paper was done in R version 3.6.0 (2019-04-26) (Team, 2019), platform: x86\_64-redhat-linux-gnu (64-bit). All code for the model and the analysis is published on <https://github.com/rivm-syso/cm-evaluation>.

<sup>a</sup> Value of the baseline parameter set; see Supplement S2 for values used in sensitivity analyses and scenarios

<sup>b</sup> Reasoned assumptions or dataset: Osiris = case notification registry; coronIT = test data; HPZone = contact tracing data; PanelClix = questionnaire in PanelClix internet panel; see Supplement for more detail

described in Section 2.2. The exact value of  $R_{ref}$  is not essential as our interest was to calculate the relative reduction in  $R$ . A value larger than one is necessary for an outbreak to enable computations. The value  $R_{ref} = 1.3$  close to one is most representative for the evaluation period and is used for all scenarios that were evaluated.

In the baseline parameter set, 10 % of the population would never test and therefore never go into isolation or quarantine, and 16 % would use the app. In the baseline simulations, these behaviours were randomly drawn for each individual, but behaviours may be clustered in social networks. We explored the effect of clustering by dividing the population into two equal parts, one with higher adherence (5 % never testing, 24 % app use) and one with lower adherence (15 % never testing, 8 % app use), and determining the effectiveness in the two subpopulations.

Sensitivity analyses were carried out with alternative parameter sets or assumptions, exploring the main uncertainties of the baseline parameters (details in Supplement S2): a lower reproduction number, a longer generation interval, less or more superspreading, less or more clustering in the tracing network, more asymptomatic infections, lower tracing probabilities, and a smaller or larger proportion of the population not adhering to anything. In addition, we considered the possibility that adherence was socially clustered, so that a mean adherence level actually results high adherence in some transmission trees, and low adherence in other transmission trees. For this, we calculated the effectiveness in low-adherence and high-adherence transmission trees.

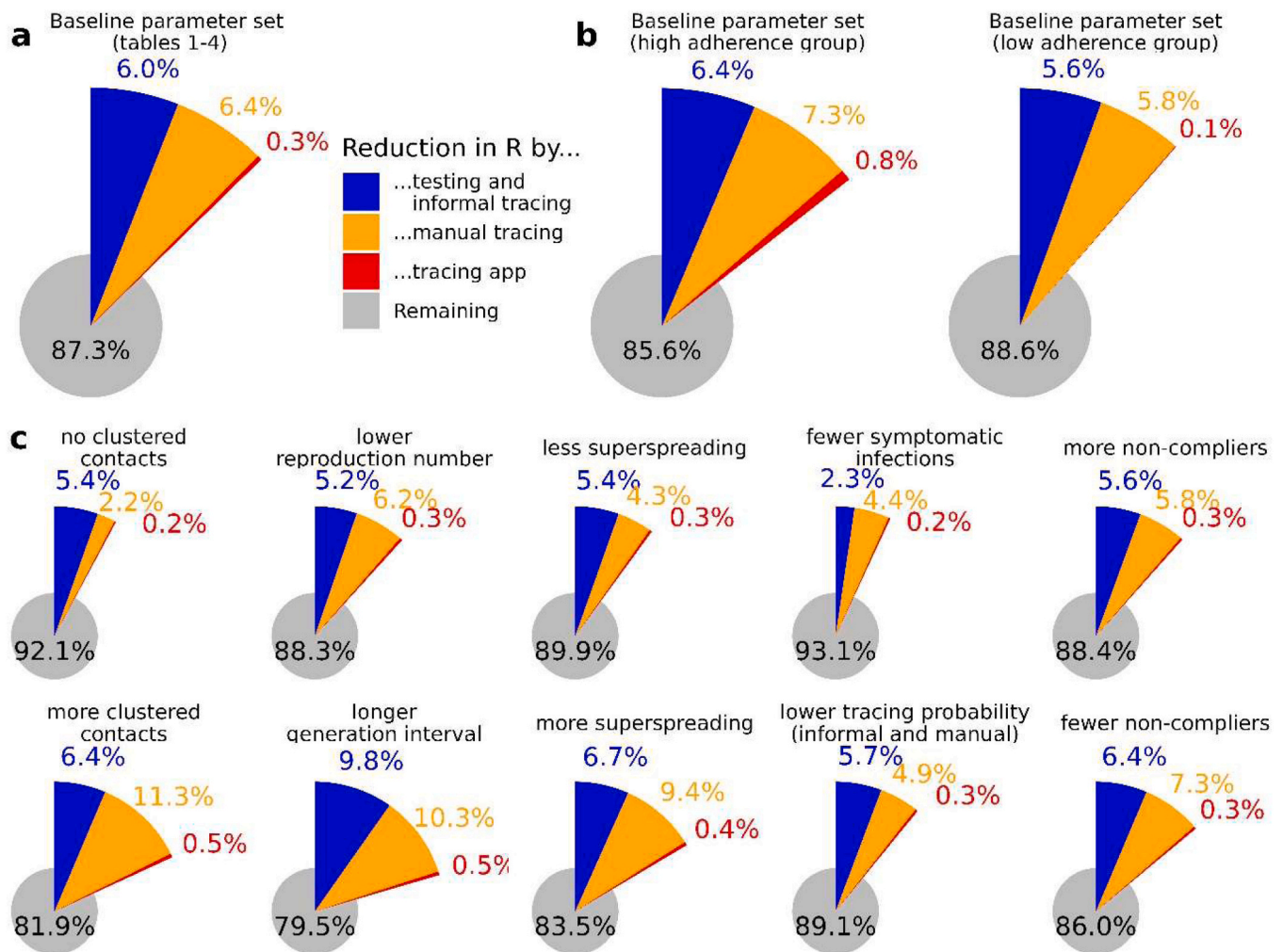
We also explored scenarios in which the contribution of the contact tracing app could be improved, if circumstances would change in the period following the study period. We evaluated several scenarios to explore the potential contribution of the contact tracing app in the “future”. First, we considered a change in policy in letting app users notify their contacts directly in the app and not through the health authorities. Second, we explored the effect of the likely reduced effectiveness of manual tracing if the lockdowns (in place during the evaluation period) would be lifted. This was expected to increase contact rates, especially with strangers, making informal and manual contact tracing less effective. In another scenario, it was even considered to completely terminate formal manual contact tracing. Finally we considered the possibility of increasing the percentage of the population using the app (by campaigns). The different scenarios are described by parameter sets for the model. They were implemented to evaluate the scenarios using the simulation model. The exact values for the parameter sets for scenarios are in Supplement S2.

## 3. Results

### 3.1. Baseline parameters and sensitivity analysis

The baseline parameter set (Tables 1–3) resulted in a total reduction of the reproduction number  $R$  by 12.7 %, which means there is a remaining 87.3 % of the reference reproduction number that is not affected by testing and tracing. The effectiveness of 12.7 % can be further specified: testing and informal tracing (notification of contactees by infected people themselves) contributed 6 %, manual tracing (notification by health authorities) another 6.4 %, and the contact tracing app only 0.3 % (Fig. 1a). This was our best estimate of the effectiveness of testing and tracing between December 2020 and March 2021 in the Netherlands. If behaviour is clustered in the population, high-adherence groups are better protected by less circulation in their social network (Fig. 1b).

In Fig. 1c the results of the sensitivity analyses of individual parameters with the most uncertainty are presented. The effectiveness of testing and informal tracing did not change very much in most analyses, and became only much less effective with more asymptomatic infections, and more effective with a longer generation interval. Manual tracing effectiveness was very sensitive to changes in the contact network (superspreading and clustering), but also to parameters directly



**Fig. 1.** Relative reduction in the reproduction number by testing and tracing, with contributions of the control levels of testing and informal tracing, manual tracing, and tracing app and the remaining percentage of the reproduction number that is not affected by testing and tracing in gray; (a) results with the baseline parameter set with parameter values reported in Tables 1–3; (b) results with the baseline parameter set where the population is subdivided into two separate groups with low and high adherence (see sections 2.4 and 2.6) (c) sensitivity analysis with each analysis adjusting exactly one parameter (see Section 2.3.5 and Supplement S2 for parameter values).

important to the test-and-trace cycle such as the proportion of asymptomatic infection, the probability to trace a contact, and the generation interval which directly affects timely notification of close contacts. Finally, the contribution of app-based tracing remained limited in all sensitivity analyses: 0.2–0.5 % reduction in  $R$ .

### 3.2. Scenarios to explore improvements to the app and its use

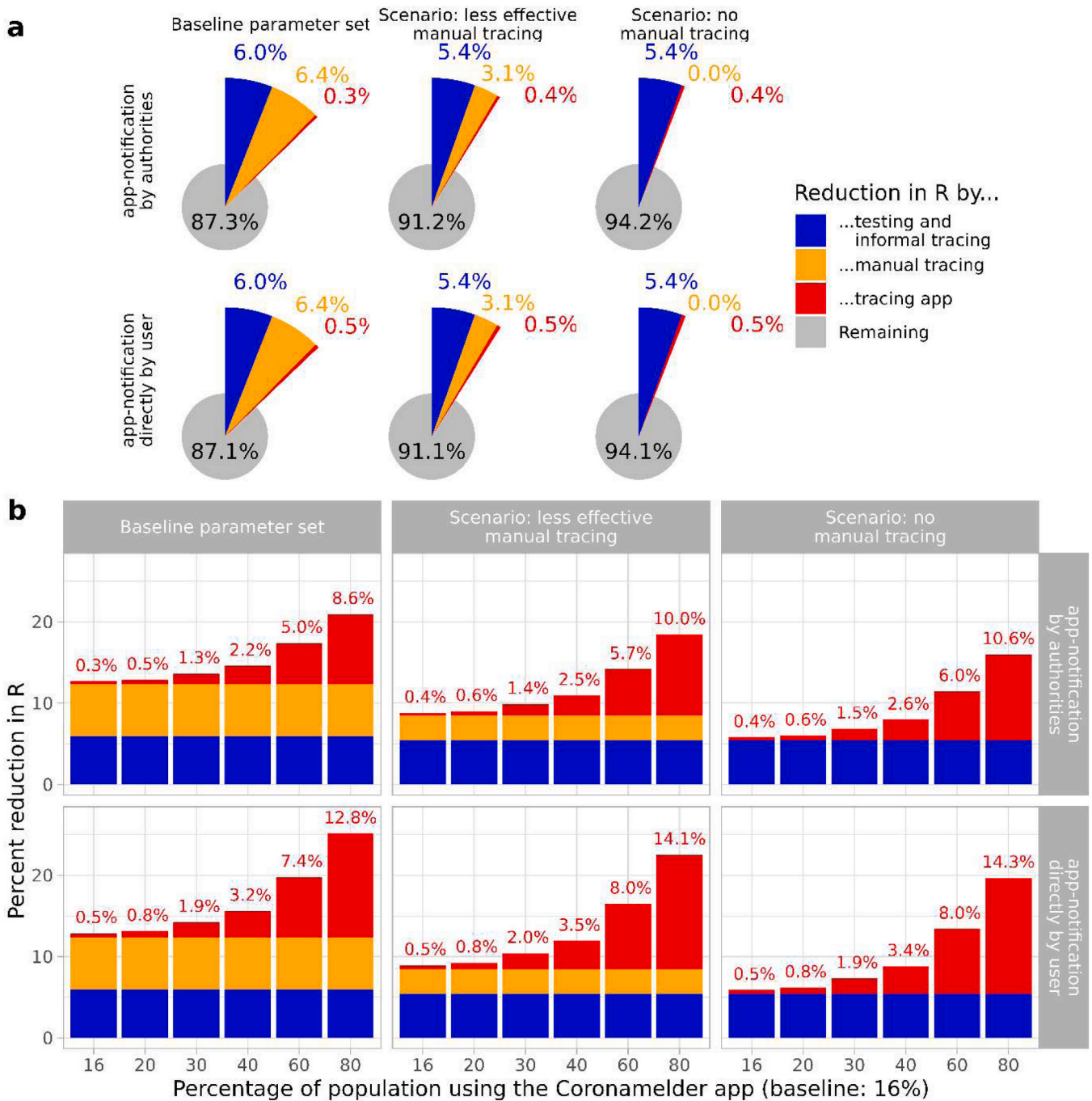
From the baseline analysis it turned out that the contact tracing app did not contribute much to transmission control. We ran a series of scenario simulations to see if this could change in other conditions. First, we looked at the possibility to change the app such that contact notification can be done by the users themselves, before they are contacted by the health authorities for manual tracing (this became possible later in 2021). This increased the reduction in  $R$  by app-based tracing from 0.3 % to 0.5 %, which in relative terms is a lot but not so much in absolute terms (Fig. 2a). Then we looked at how the app would work if lockdown measures were lifted, rendering manual tracing less effective because more contacts are made with strangers, and how it would work when manual tracing would be abandoned at all. It turned out that the reduced contribution of manual tracing was not expected to be compensated by the contact tracing app (Fig. 2a).

Finally, we explored how much more effective the contact tracing

app would become with more users (Fig. 2b), in combination with direct notification by the users and a reduction in manual tracing effectiveness or abandonment of manual tracing. More users made the contact tracing app more effective, but to compensate for the expected lower effectiveness of manual tracing in a scenario without lockdown, at least 40 % of population would need to use the contact tracing app instead of the 16 % in the baseline scenario. In that case, the benefit of notification by users themselves also increases, and would contribute 1 % of the total reduction of 12.5 % (2.5 % vs 3.5 % - top vs bottom row in Fig. 2b, middle column). In the extreme case that 80 % of the population would use the contact tracing app, the app could contribute 14.1 % of the reduction in  $R$  (Fig. 2b middle column, bottom row).

## 4. Discussion and conclusion

We evaluated the epidemiological effectiveness of the Dutch contact tracing app CoronaMelder in the period from December 2020 to March 2021. We used a simulation model that was informed by testing and tracing data collected in the Netherlands during this study period. We conclude that the contribution of the contact tracing app to control the epidemic was very small, although some protection may have been provided to social groups with more app use. This conclusion was not very sensitive to the major parameter uncertainties. The app could have



**Fig. 2.** Analysis for scenarios for the effectiveness of testing, manual tracing, and digital contact tracing in the base setting and a society with (little) restrictions, both with and without formal manual tracing (a) Reduction in the reproduction number, in the top row the app notification is by authorities while in the bottom row the app notification is by the user him/herself (24 vs 8 hours) (b) For the same six parameter sets the effect of a higher percentage app users on the reduction of the reproduction number.

contributed more if more people would have used it, and if notification of contacts could have been done directly by the user, without involving the health authorities. If formal manual tracing would have become less effective or even abandoned, this would not have been compensated by the app (Fig. 2a: the effectiveness of the app does not increase enough to compensate the decrease in effectiveness of manual tracing).

The results in our study are at the lower end of the spectrum of effectiveness estimates of contact tracing apps (Jenniskens et al., 2021). It is not straightforward to compare studies, as there are many complexities that play a role (Kretzschmar et al., 2022). Some modelling studies show the effectiveness of the digital contact tracing app in a

general setting as a proof of concept, while other studies, like ours, are tailored to a specific country and period (and corresponding policies in place). In particular, the time intervals (Table 3) that were estimated from the data during the evaluation period were relatively long compared to most of the intervals used in literature. It is known that longer time delays in test-and-trace methods can greatly reduce the effectiveness of tracing (Fraser et al., 2004; Klinkenberg et al., 2006; Kretzschmar et al., 2020). In addition, the percentage of app users was low in the Netherlands as compared to the values used in other studies, e.g. Kretzschmar et al. (2020) consider scenarios with 60 %, 80 % and 100 % app usage, Kucharski et al. (2020) estimated an app-usage of 53

% and Ferretti et al. (2020) estimated 28 % app users in the UK. Simulations with more users and shorter delays showed the potential benefit of a contact tracing app (Fig. 2), which should therefore certainly not be dismissed as a tool in controlling a future epidemic.

Many models for contact tracing have been built, also for evaluation of digital contact tracing during the SARS-CoV-2 pandemic (e.g. (Bradshaw et al., 2021; Cencetti et al., 2021; Kretzschmar et al., 2020; Kucharski et al., 2020; Quilty et al., 2021; Scott et al., 2021)), but our model was unique on three points. First of all, we were able to obtain realistic values for most parameters based on a large variety of data sources of high quality, such as large epidemiological surveillance systems and anonymous questionnaire surveys about behaviour in representative samples of the Dutch population. An important consequence is that our results were a quantitative estimate of the effectiveness of the tracing app in the Netherlands, which could be easily interpreted by policy makers. In fact, the evaluation of the CoronaMelder app (of which the epidemiological evaluation by this model was one part) was followed first by some improvements such as direct notifications by the user, but ultimately by termination of its use (Ministry of Health, 2022).

Second, we made an explicit distinction between formal manual tracing and informal tracing by cases notifying their contacts themselves after receiving a positive test result. Informal tracing was observed in behavioural surveys, and is fast and cheap compared to manual tracing. A proper evaluation of tracing programmes should take this form of tracing into account, because it reduces the additional benefit of manual tracing. On the other hand, manual tracing may reach more contactees. Moreover, in manual tracing, instructions are provided by official health authorities, which may improve the effectiveness of quarantine. In the evaluated context in the Netherlands, manual tracing was estimated to add 6.4 % reduction in  $R$  to the 6 % from testing and informal tracing (Fig. 1a).

Third, we included the possibility of tracing of contacts that were infected by someone else in the social network, known to be important for the effectiveness of tracing (House and Keeling, 2010; Kretzschmar et al., 2022). This was done by adding contacts to the simulated transmission trees through which tracing could take place. The extent to which these extra contacts are added makes a large difference to the effectiveness of tracing (sensitivity analysis in Fig. 1c), which is easily understood as these contacts can bypass the transmission links and lead to quarantine just after or even prior to infection. In our model, we included these contacts so that triangles are created in the original transmission tree, just like triangles are part of clustered networks (Kiss et al., 2017; Newman, 2003). In clustered networks, the clustering coefficient  $c$  determines what proportion of triplets form triangles in the network. A topic for further study is how clustering among infected individuals in a transmission tree (as in our model) relates to clustering in an underlying social network.

Next to the strengths of our model, we would like to discuss some limitations and how they may have affected the results. Delays were fixed in the model. Including variation in delays in the model would likely increase the effectiveness of testing and tracing, including the tracing app, as the variation would mean that a part of the testing and tracing process would have shorter delays leading to more cases that go into isolation at an earlier time (while longer delays result in the same onward transmission). As many time intervals are involved in the process, we expect only moderate increase in the effectiveness. Another aspect that we did not take into account is population age structure or other population heterogeneity such as socioeconomic status. It is plausible that app use and adherence to measures may correlate to age or social norms. They may lead to clustering in app use and adherence and can lead to higher effectiveness in these groups. We investigated the impact of high and low adherence groups in the sensitivity analysis; see Fig. 1b. It would be interesting to further investigate clustering in future research. Behaviour in the form of app use and adherence to measures was modelled as fixed throughout the evaluation period. One can imagine changes in behaviour over time e.g. due to fatigue or due to external

factors such as changes in the epidemic or reporting in the media. It is hard to speculate in what way the effectiveness of the tracing app would change if we would incorporate behaviour change, and therefore interesting to investigate how behaviour change could be taken into account in infectious disease models.

### CRedit authorship contribution statement

**KY Leung:** Conceptualization, Methodology, Software, Analysis, Writing – Original Draft, **E Metting:** Conceptualization, Methodology, Data resources, **W Ebbers:** Conceptualization, Methodology, Data resources, **I Veldhuijzen:** Methodology, Data resources, **SP Andeweg:** Methodology, Data resources, **G Luijben:** Methodology, Data resources, **M de Bruin:** Methodology, Data resources, **J Wallinga:** Conceptualization, **D Klinkenberg:** Conceptualization, Methodology, Software, Analysis, Writing – Original Draft. All authors have contributed to reviewing and editing of the manuscript.

### Declaration of Competing Interest

None.

### Data availability

<https://github.com/rivm-syso/cm-evaluation>.

### References

- Ashraf, B.N., 2020. Economic impact of government interventions during the COVID-19 pandemic: international evidence from financial markets. *J. Behav. Exp. Financ.* 27, 100371 <https://doi.org/10.1016/j.jbef.2020.100371>.
- Backer, J.A., Klinkenberg, D., Wallinga, J., 2020. Incubation period of 2019 novel coronavirus (2019-nCoV) infections among travellers from Wuhan, China, 20-28 January 2020. *Eur. Surveill.* 25 <https://doi.org/10.2807/1560-7917.ES.2020.25.5.2000062>.
- Bradshaw, W.J., Alley, E.C., Huggins, J.H., Lloyd, A.L., Esvelt, K.M., 2021. Bidirectional contact tracing could dramatically improve COVID-19 control. *Nat. Commun.* 12, 232 <https://doi.org/10.1038/s41467-020-20325-7>.
- Braithwaite, I., Callender, T., Bullock, M., Aldridge, R.W., 2020. Automated and partly automated contact tracing: a systematic review to inform the control of COVID-19. *Lancet Digit Health* 2, e607–e621. [https://doi.org/10.1016/S2589-7500\(20\)30184-9](https://doi.org/10.1016/S2589-7500(20)30184-9).
- Buitrago-Garcia, D., Egli-Gany, D., Counotte, M.J., Hossmann, S., Imeri, H., Ipekci, A.M., Salanti, G., Low, N., 2020. Occurrence and transmission potential of asymptomatic and presymptomatic SARS-CoV-2 infections: A living systematic review and meta-analysis. *PLoS Med.* 17, e1003346 <https://doi.org/10.1371/journal.pmed.1003346>.
- Cencetti, G., Santin, G., Longa, A., Pigani, E., Barrat, A., Cattuto, C., Lehmann, S., Salathe, M., Lepri, B., 2021. Digital proximity tracing on empirical contact networks for pandemic control. *Nat. Commun.* 12, 1655 <https://doi.org/10.1038/s41467-021-21809-w>.
- Cheng, H.-Y.J., Liu, S.-W., Ng, D.-P., Huang, T.-C., Lin, H.-H., W.-T., 2020. High transmissibility of COVID-19 near symptom onset. *medRxiv*. <https://doi.org/10.1101/2020.03.18.20034561>.
- Eames, K.T., Keeling, M.J., 2003. Contact tracing and disease control. *Proc. Biol. Sci.* 270, 2565–2571. <https://doi.org/10.1098/rspb.2003.2554>.
- Ebbers, W., 2021. Wat volgt er na een melding van CoronaMelder? Eindrapportage van de evaluatie van 'CoronaMelder-app meldingen van een kans op besmetting', een kwantitatief onderzoek. Erasmus University Rotterdam.
- Ferrari, A., Santus, E., Cirillo, D., Ponce-de-Leon, M., Marino, N., Ferretti, M.T., Santuccioni Chadha, A., Mavridis, N., Valencia, A., 2021. Simulating SARS-CoV-2 epidemics by region-specific variables and modeling contact tracing app containment. *NPJ Digit Med.* 4, 9 <https://doi.org/10.1038/s41746-020-00374-4>.
- Ferretti, L., Wymant, C., Kendall, M., Zhao, L., Nurtay, A., Abeler-Dorner, L., Parker, M., Bonsall, D., Fraser, C., 2020. Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing. *Science* 368. <https://doi.org/10.1126/science.abb6936>.
- Fraser, C., Riley, S., Anderson, R.M., Ferguson, N.M., 2004. Factors that make an infectious disease outbreak controllable. *Proc. Natl. Acad. Sci. USA* 101, 6146–6151. <https://doi.org/10.1073/pnas.0307506101>.
- Ge, Y., Chen, Z., Handel, A., Martinez, L., Xiao, Q., Li, C., Chen, E., Pan, J., Li, Y., Ling, F., Shen, Y., 2021. The impact of social distancing, contact tracing, and case isolation interventions to suppress the COVID-19 epidemic: a modeling study. *Epidemics* 36, 100483. <https://doi.org/10.1016/j.epidem.2021.100483>.
- Hellewell, J., Abbott, S., Gimma, A., Bosse, N.I., Jarvis, C.I., Russell, T.W., Munday, J.D., Kucharski, A.J., Edmunds, W.J., Centre for the Mathematical Modelling of Infectious Diseases, Funk, C.-W.G., Eggo, S., R.M., 2020. Feasibility of controlling

- COVID-19 outbreaks by isolation of cases and contacts. *Lancet Glob. Health* 8, e488–e496. [https://doi.org/10.1016/S2214-109X\(20\)30074-7](https://doi.org/10.1016/S2214-109X(20)30074-7).
- House, T., Keeling, M.J., 2010. The impact of contact tracing in clustered populations. *PLoS Comput. Biol.* 6, e1000721 <https://doi.org/10.1371/journal.pcbi.1000721>.
- Jenniskens, K., Bootsma, M.C.J., Damen, J., Oerbetke, M.S., Vernooij, R.W.M., Spijker, R., Moons, K.G.M., Kretzschmar, M.E.E., Hooft, L., 2021. Effectiveness of contact tracing apps for SARS-CoV-2: a rapid systematic review. *BMJ Open* 11, e050519. <https://doi.org/10.1136/bmjopen-2021-050519>.
- Kirkegaard, J.B., Sneppen, K., 2021. Superspreading quantified from bursty epidemic trajectories. *Sci. Rep.* 11, 24124 <https://doi.org/10.1038/s41598-021-03126-w>.
- Kiss, I.Z., Miller, J.C., Simon, P.L., 2017. *Mathematics of epidemics on networks: from exact to approximate models*. Springer, Cham.
- Kissler, S.M.K., Tang, P., Conlan, M., Gog, J.R., A.J.K., 2020. Sparking "The BBC Four Pandemic": leveraging citizen science and mobile phones to model the spread of disease. *bioRxiv*. <https://doi.org/10.1101/479154>.
- Klinkenberg, D., Fraser, C., Heesterbeek, H., 2006. The effectiveness of contact tracing in emerging epidemics. *PLoS One* 1, e12. <https://doi.org/10.1371/journal.pone.0000012>.
- Klinkenberg, D., Leung, K.Y., Wallinga, J., 2021. The CoronaMelder app - a model study into effectiveness. Digital contact tracing in the fight against COVID-19. Netherlands Institute for Public Health and the Environment. <https://doi.org/10.21945/RIVM-2021-0092>.
- Kretzschmar, M.E., Ashby, B., Fearon, E., Overton, C.E., Panovska-Griffiths, J., Pellis, L., Quaipe, M., Rozhnova, G., Scarabel, F., Stage, H.B., Swallow, B., Thompson, R.N., Tildesley, M.J., Villela, D., 2022. Challenges for modelling interventions for future pandemics. *Epidemics* 38, 100546. <https://doi.org/10.1016/j.epidem.2022.100546>.
- Kretzschmar, M.E., Rozhnova, G., Bootsma, M.C.J., van Boven, M., van de Wijgert, J., Bonten, M.J.M., 2020. Impact of delays on effectiveness of contact tracing strategies for COVID-19: a modelling study. *Lancet Public Health* 5, e452–e459. [https://doi.org/10.1016/S2468-2667\(20\)30157-2](https://doi.org/10.1016/S2468-2667(20)30157-2).
- Kretzschmar, M.E., Rozhnova, G., van Boven, M., 2021. Isolation and contact tracing can tip the scale to containment of COVID-19 in populations with social distancing. *Front. Phys.* 8 <https://doi.org/10.3389/fphy.2020.622485>.
- Kucharski, A.J., Klepac, P., Conlan, A.J.K., Kissler, S.M., Tang, M.L., Fry, H., Gog, J.R., Edmunds, W.J., group, C.C.-w., 2020. Effectiveness of isolation, testing, contact tracing, and physical distancing on reducing transmission of SARS-CoV-2 in different settings: a mathematical modelling study. *Lancet Infect. Dis.* 20, 1151–1160. [https://doi.org/10.1016/S1473-3099\(20\)30457-6](https://doi.org/10.1016/S1473-3099(20)30457-6).
- Kucharski, A.J., Wenham, C., Brownlee, P., Racon, L., Widmer, N., Eames, K.T.D., Conlan, A.J.K., 2018. Structure and consistency of self-reported social contact networks in British secondary schools. *PLoS One* 13, e0200090. <https://doi.org/10.1371/journal.pone.0200090>.
- Kurita, J., Sugawara, T., Ohkusa, Y., 2021. Estimated effectiveness of school closure and voluntary event cancellation as COVID-19 countermeasures in Japan. *J. Infect. Chemother.* 27, 62–64. <https://doi.org/10.1016/j.jiac.2020.08.012>.
- McDonald, S.A., Miura, F., Vos, E.R.A., van Boven, M., de Melker, H.E., van der Klis, F.R.M., van Binnendijk, R.S., den Hartog, G., Wallinga, J., 2021. Estimating the asymptomatic proportion of SARS-CoV-2 infection in the general population: Analysis of nationwide serosurvey data in the Netherlands. *Eur. J. Epidemiol.* 36, 735–739. <https://doi.org/10.1007/s10654-021-00768-y>.
- Ministry of Health, W.a.S., 2020. Veldtest bluetooth validatie COVID-19 notificatie app. (<https://www.rijksoverheid.nl/documenten/rapporten/2020/06/08/veldtest-bluetooth-validatie-covid-19-notificatie-app>), accessed 2023 2 May.
- Ministry of Health, W.a.S., 2021a. CoronaMelder herctest bluetooth validatie beoordeling AEN v2 API. (<https://www.rijksoverheid.nl/ministeries/ministerie-van-volksgezondheid-welzijn-en-sport/documenten/publicaties/2021/05/17/coronamelder-herctest-bluetooth-validatie-april-2021>), accessed 2023 2 May.
- Ministry of Health, W.a.S., 2021b. CoronaMelder statistics. (<https://github.com/minvws/nl-covid19-notification-app-statistics>), accessed 2023 2 May.
- Ministry of Health, W.a.S., 2022. CoronaMelder. (<https://www.rijksoverheid.nl/onderwerpen/coronavirus-covid-19/coronamelder>), 2 May 2023.
- Muller, J., Kretzschmar, M., Dietz, K., 2000. Contact tracing in stochastic and deterministic epidemic models. *Math. Biosci.* 164, 39–64. [https://doi.org/10.1016/S0025-5564\(99\)00061-9](https://doi.org/10.1016/S0025-5564(99)00061-9).
- Newman, M.E.J., 2003. Properties of highly clustered networks. *Phys. Rev. E* 68, 026121.
- Nuzzo, A., Tan, C.O., Raskar, R., DeSimone, D.C., Kapa, S., Gupta, R., 2020. Universal shelter-in-place versus advanced automated contact tracing and targeted isolation: a case for 21st-century technologies for SARS-CoV-2 and future pandemics. *Mayo Clin. Proc.* 95, 1898–1905. <https://doi.org/10.1016/j.mayocp.2020.06.027>.
- Quilty, B.J., Clifford, S., Hellewell, J., Russell, T.W., Kucharski, A.J., Flasche, S., Edmunds, W.J., Centre for the Mathematical Modelling of Infectious Diseases, C.-w.g., 2021. Quarantine and testing strategies in contact tracing for SARS-CoV-2: a modelling study. *Lancet Public Health* 6, e175–e183. [https://doi.org/10.1016/S2468-2667\(20\)30308-X](https://doi.org/10.1016/S2468-2667(20)30308-X).
- RIVM, 2021. Resultaten onderzoek gedragsregels en welbevinden [results research behavioural rules and wellbeing]. (<https://www.rivm.nl/gedragsonderzoek/maatregelen-welbevinden>), accessed 2023 2 May.
- Scott, N., Palmer, A., Delpont, D., Abeyuriya, R., Stuart, R.M., Kerr, C.C., Mistry, D., Klein, D.J., Sacks-Davis, R., Heath, K., Hainsworth, S.W., Pedrana, A., Stooze, M., Wilson, D., Hellard, M.E., 2021. Modelling the impact of relaxing COVID-19 control measures during a period of low viral transmission. *Med. J. Aust.* 214, 79–83. <https://doi.org/10.5694/mja2.50845>.
- Susswein, Z.B., S., 2020. Characterizing superspreading of SARS-CoV-2: from mechanism to measurement. *medRxiv*. <https://doi.org/10.1101/2020.12.08.20246082>.
- Team, R.C., 2019. R: A language and environment for statistical computing. R foundation for statistical computing, Vienna, Austria.
- Van den Boom, W.M.E., Van der Vliet, N., Elberse, J., Van Dijken, S., Van Dijk, M., Euser, S., Derks, M., Leurs, M., Albers, C., Sanderma, R., De Bruin, M., 2022. Mental health, loneliness and life satisfaction in the dynamic context of COVID-19 pandemic policies and healthcare burden: a two-year cohort study in the Netherlands. *OSF Prepr.* <https://doi.org/10.31219/osf.io/qr8fv>.
- van der Laan, N., van der Waal, N., de Wit, J., 2021. Endreport CoronaMelder Evaluation - Survey LISS panel wave 3. Tilburg University.
- Ward, M., Brandsema, P., van Straten, E., Bosman, A., 2005. Electronic reporting improves timeliness and completeness of infectious disease notification, The Netherlands, 2003. *Eur. Surveill.* 10, 7–8. <https://doi.org/10.2807/10.01.00513-en>.