An imperfect tool: COVID-19 'test & trace' success relies on minimising the impact of false negatives and continuation of physical distancing.

Emma L. Davis[1,+,*], Tim C. D. Lucas[1,+], Anna Borlase[1], Timothy M Pollington[1,3], Sam Abbott[2], Diepreye Ayabina[1], Thomas Crellen[1], Joel Hellewell[2], Li Pi[1], CMMID COVID-19 working group[6], Graham F. Medley[4], T. Déirdre Hollingsworth[1,†], and Petra Klepac[2,5,†]

[1]Big Data Institute, University of Oxford, UK.
[2]Department of Infectious Disease Epidemiology, London School of Hygiene and Tropical Medicine, UK.
[3]MathSys CDT, University of Warwick, UK.
[4]Centre for Mathematical Modelling of Infectious Disease & Department of Global Health and Development, London School of Hygiene and Tropical Medicine, London, UK.
[5]Department for Applied Mathematics and Theoretical Physics, University of Cambridge, UK.
[6]Full list of members in supplementary material.
[*]Corresponding author: Emma.Davis@bdi.ox.ac.uk
[+,†]Contributed equally

Abstract

The increasingly evident role of asymptomatic and pre-symptomatic transmission means testing is central to COVID-19 control, but test sensitivity estimates are low (around 65%). We extend an existing branching process contact tracing model, adding diagnostic testing and refining parameter estimates. Poor test sensitivity potentially reduces the efficacy of contact tracing, due to false-negative results impacting quarantine. We show that, counter-intuitively, faster testing could also reduce operational test sensitivity, exacerbating this effect. If sensitivity-based risks are mitigated, we find that contact tracing can facilitate control, but small changes in the population reproduction number (1·3 to 1·5) could impact contact tracing feasibility.

Main

In December 2019, SARS-CoV-2, a novel coronavirus strain, was detected in Hubei Province, China [1]. By 31st January 2020 the first UK cases of COVID-19, the disease caused by SARS-CoV-2, were confirmed [2]. Initial modelling studies indicated that fast and effective contact tracing could contain the UK outbreak in most settings [3, 4]. However, by 20th March there were almost 4,000 confirmed cases nationwide [5], at which point the UK Government halted national contact tracing and scaled up physical distancing measures, including the closure of schools and social venues, extending to heightened restrictions on non-essential travel, outdoor activities and between-household social mixing [6]. Similar patterns occurred in other countries [7, 8].

By early May 2020 these measures were estimated to have reduced the effective reproduction number, R, from 2·6 to 0·62 [9, 10] and so from 12th–13th May in England some limitations on outdoor exercise were lifted and workers encouraged to return to work if they could maintain physical distancing [11].

Capacity for diagnostic testing in the UK, as in other countries, has been escalated over recent months, with capacity reaching over 300,000 tests a day by the end of June (`https://coronavirus.data.gov.uk/testing`). Currently, testing of asymptomatic individuals is limited to staff and patients in NHS and social care facilities [12], but on the 28th of May the UK Government rolled out the initial stages of their 'Test & Trace' contact tracing programme to the general population, which aims to follow chains of transmission and use isolation to prevent onward spread. Since the beginning of 'Test & Trace' over 3 million people were tested by the end of July, just under 50,000 of which were positive [13].

Crucially, the current strategy only tests symptomatic contacts and notifies individuals that they no longer need to isolate following a negative test, which comes back in 24 or 48 hours depending on type of test [13]. However, there are critical limitations to the diagnostic test, with poor sensitivity (current estimates imply close to 65% [14, 15]), especially in community-based settings, leading to high false negative rates which are exacerbated by high variability in symptom severity [15]. Infectious individuals who test falsely negative may prematurely resume their normal activities, contributing to ongoing chains of transmission: reliance on testing of contacts will always reduce the effectiveness of contact tracing, potentially substantially [16].

Imperfect adherence and the innate difficulties in identifying contacts will pose challenges for 'Test & Trace', particularly in crowded urban settings [17]. Therefore, evaluating both the limitations of contact tracing and how to maximise its effectiveness could be crucial in preventing an exponential rise in cases, which might see contact tracing capacity rapidly exceeded and stricter physical distancing measures required [18].

As our knowledge of the transmission dynamics of SARS-CoV-2 grows, extending Hellewell et al.'s [3] UK-focused contact tracing study with new insights could inform this 'Test & Trace' strategy. The key conclusion of the initial study was that highly effective contact tracing would be sufficient to control an initial outbreak of COVID-19 in the UK, however substantial new evidence supports much higher pre- and asymptomatic transmission rates than had initially been considered [19, 20, 21]. The focus on rapid testing in the UK contact tracing programme also requires a detailed assessment of the associated trade-offs through mechanistic modelling of the testing process. Up-to-date modelling studies are therefore needed to investigate the feasibility of contact tracing and the conditions under which it is effective.

We use improved incubation period and serial interval estimates [22, 23, 24], consider imperfect self-reporting and tracing rates and simulate the use of diagnostic tests for both detection and tracing of asymptomatic infection chains. We also simulate decision-making regarding quarantine procedures for traced individuals and explore the trade-offs introduced by poor test sensitivity, particularly when negative test results are used to advise individuals to cease self-isolation.

Results

Testing

By comparing the time individuals are tested after exposure in the model to temporal estimates of PCR test sensitivity [15] we see that an average of 65% sensitivity is a relatively realistic expectation if testing is conducted 2 or more days post-isolation (Figure 1a). Test sensitivity is expected to peak at just over 75% 8 days post-exposure, but the majority of testing in the model occurs between 4 and 7 days post-exposure, where sensitivity estimates have more variance.

In the case of a 2 day testing delay, 29·3% of cases are tested before 5 days post-exposure, at which point sensitivity estimates are around 62%, and only

10·2% of cases are tested before 4 days post-exposure. The time-weighted average of expected test sensitivity for all cases tested in the model with a 2 day delay is 68·0%. For a 4 day test delay, less than 1% of cases are tested before day 5 post-exposure, but just under 10% are tested after day 14, when test sensitivity drops to 61%.

However, in the case of immediate testing, 65% sensitivity could be a substantial over-estimate, with 64·5% of cases being tested before 5 days post-exposure and 48·9% before 4 days, meaning the test could be less than 33% sensitive for around half of all cases tested.

Even if test sensitivity was constant, the timing of testing and quarantine duration has an impact on the risk of a large outbreak (at least 2,000 cases) and can undermine the positive impact of testing within a contact tracing programme. The probability of a large outbreak occurring is greater with an assumed test sensitivity of 65% compared to scenarios where no testing was carried out at all if testing is rapid and results in an immediate return to normal behaviour (Figure 1b, upper left panel). This result was observed across all contact tracing coverage rates. The deleterious effect of releasing false negative cases is mitigated by using a precautionary seven-day quarantine period, which reduced the risk of a large outbreak 11·7% to 3·9% for $R_S = 1{\cdot}3$ with 60% contact tracing (Figure 1b).

A two day delay in carrying out the tests also led to a decrease in the probability of a large outbreak, from 11·7% to 6·3% for R_S of 1·3. Combining the two-day delay in testing and the seven-day precautionary quarantine reduced the risk of a large outbreak further, from 11·7% to 3·4% for $R_S = 1{\cdot}3$ and 60% contact tracing.

In the case of instant testing and an immediate end to quarantine if the test is negative, there was a comparatively small benefit from scaling up of contact tracing coverage from 0% to 100%, implying that much of the potential positive impact of contact tracing could be lost if such an approach were taken.

Whilst a test with 65% sensitivity and no minimum quarantine period can reduce the benefits of contact tracing, if a test were to be 95% sensitive, this would improve the outcome compared to no testing in all scenarios. However, with a two-day test delay and seven-day precautionary quarantine a 65% sensitive test is almost as effective in reducing transmission as a 95% sensitive test due to this strategy ensuring quarantine of all cases during peak transmission periods irrespective of test result.

Case detection

Even with perfect contact tracing and employing good diagnostic practices (100% of contacts traced in 24 hours and a minimum quarantine period of 7 days), a large proportion of cases are likely to go unobserved (Figure 2). High levels of symptomatic self-reporting to the tracing programme and improved test sensitivity can increase case detection: 95% sensitivity and 100% self-reporting gives an increase from 30·5% to 73·9% compared to 65% sensitivity and 50% self-reporting (both for $R_S = 1.3$). However, this still results in 26·1% of cases being missed, hence detecting every case is essentially infeasible.

Super-spreading events

Every missed case is a potential new chain of transmission and, given the high heterogeneity in the secondary case distribution, characterised by dispersion parameter k, there is a substantial risk of super-spreading events. Considering a scenario with poor adherence to self-reporting guidelines and where one missed case leads to a cluster of either 5 or 100 new cases, we assume observation of the outbreak only occurs when the first case is hospitalised, after which contact tracing may be initiated (Figure 3a and b).

For a cluster of 5 new cases the median total unobserved outbreak size before the first case is hospitalised is 13 cases for $R_S = 1{\cdot}3$ and 16 cases for $R_S = 1{\cdot}5$, which translates to 8·0% and 36·9% probability of a large outbreak respectively if 60% contact tracing can be implemented (Figure 3c). For a cluster of 100 new cases the median is 219 for $R_S = 1{\cdot}3$ and 238 for $R_S = 1{\cdot}5$, translating to 36·3% and 84·9% probability of a large outbreak with 60% contact tracing. This emphasises the importance of maintaining physical distancing measures that restrict the attendance of indoor social gatherings to avoid super-spreading events which could rapidly escalate.

Additional observations

More generally, the probability of a large outbreak given the current outbreak size (Figure 3c) could be used to assess at what point during an epidemic contact tracing would be unable to control transmission, as well as to inform targets for coverage and speed. In our model both the time taken to trace contacts and the proportion of contacts traced had effects on the risk of a large outbreak, although increasing tracing speed may have the counter-effect

of reducing average test sensitivity due to impact on test timing with respect to exposure.

We also found that higher contact tracing coverage results in a lower overall number of individuals which are traced, tested and quarantined, due to the lower outbreak size. This means that achieving greater efficacy in tracing will ultimately require fewer resources. However, these resources are likely to be needed in a more condensed period of time.

Conclusions

Our results show that with a test sensitivity of 65%, fast testing which recommends infected but false-negative individuals to cease quarantine could be counter-productive, undermining contact tracing efforts, and may be worse than not testing. However the impact of low test sensitivity could be mitigated by applying a minimum quarantine period to all traced contacts and using positive tests to prompt further contact tracing. This would allow negative individuals to leave quarantine comparatively early, but not immediately upon receipt of test result. Simply slowing down the decision-making process, so any false negative tests occur later in the infectious period, will also reduce the amount of transmission caused by premature cessation of quarantine and potentially increase likelihood of a more accurate test result [15]. Control policies in some countries are being designed to account for the high proportion of false negative individuals: for instance Greece requires negative-testing international arrivals to self-quarantine for seven days [25]; in Singapore two negative tests 24 hours apart are required to release from quarantine [26].

We show that even a test with low (65%) sensitivity can improve contact tracing outcomes if the impact of false negative cases can be limited by employing appropriate precautionary measures. This effect is seen because testing can bridge asymptomatic links in transmission chains that would otherwise have been missed, although there is some uncertainty surrounding the infectiousness of asymptomatic individuals [23]. Nonetheless, this benefit is only possible, provided testing is applied to all contacts, not just those displaying symptoms as was the initial UK policy.

Testing asymptomatic contacts would require more testing and resources, as well as potentially testing individuals earlier in their infectious period, before symptom onset. Earlier testing increases the impact of immediate quarantine cessation for false negative cases, so this would require a minimum

quarantine period. Despite these considerations, if very good contact tracing can be implemented from the beginning of the outbreak then fewer total resources will be required because of a smaller final outbreak size, meaning the key factor for feasibility will be time-limited resource access.

We demonstrated that small increases in the reproduction number under physical distancing measures, R_S, has a large impact on the feasibility of contact tracing. We only consider values of R_S up to 1·5, which is still substantially lower than estimates of R_0 in the absence of any interventions ($R_0 \approx 2{\cdot}7$ [27]), but may be achievable with partial interventions. Our estimates of R_S reflect a decrease in social contacts of almost 50% but even 60% coverage and a one day trace time is insufficient to negate the risk of a large outbreak. This reiterates that physical distancing is still critical, even with highly effective contact tracing, and that contact tracing will likely be insufficient to allow a complete return to normal life without additional measures, such as an effective vaccine.

In addition to general physical distancing, the risk posed by a single large super-spreading event means that relaxing restrictions on large gatherings, particularly indoors, could lead to a rise in case numbers, especially in communities where self-reporting rates are expected to be poor. Even with very low $R_S = 1{\cdot}1$, a local cluster of 100 unobserved cases could approximately double in size before being detected.

However, we found that the risk of a large outbreak ($\geq$2,000 cases) was relatively low for $R_S = 1{\cdot}1$ no matter what the contact tracing and testing strategy. What is of note to national governments who are exiting extreme social distancing is that a dramatic change in the dynamics occurs in the small absolute increase of R_S to 1·3. At $R_S = 1{\cdot}1$ with a poorly resourced or ineffective contact tracing system the probability of a large outbreak is roughly 1%. However when $R_S \geq 1.3$ then an ineffective contact tracing system becomes noticeable, at which stage it is too late to act.

A number of our assumptions may cause our results to appear unduly optimistic. For example, we model a scenario with very low initial case numbers and assume that tracing can be initiated before test results are received, and that contacts of up to 3 days pre-onset are traced. This means there is potentially an increased requirement for maintaining physical distancing measures, even if contact tracing is deployed at high coverage nationwide.

We also consider the test to have a fixed sensitivity over the course of infection, whereas previous studies show that testing too early or late after exposure can dramatically increase false negative rates [15]. However, these

temporal estimates ignore variation in the incubation period, assuming a fixed onset time of 5 days. Additionally, high between-person variance has been observed in the natural history of infection [23]. It is therefore unclear what is driving these temporal changes in sensitivity or whether this temporal profile makes sense on an individual basis.

Furthermore there were worrying trends in adherence to movement restrictions towards the end of "lockdown", suggesting that recommended quarantine through the 'Test & Trace' programme may also be affected; an unpublished study of 90,000 adults across the UK in the two weeks up to 25th May found that lockdown adherence may have dropped to 50% [28]. Our assumption of 90% untraced symptomatic individuals self-isolating is therefore at the upper end of realistic, although symptomatic individuals will perhaps be more cautious or less mobile due to the burden of symptoms. However, this could also have repercussions on assuming that contact-traced individuals will self-isolate when asked to do so, particularly if asymptomatic.

Modelling studies in other countries have proposed combinations of contact tracing and population-level mitigation strategies [29] and a recent UK study puts R_S in the range of 1–1·6 for a combination of school closures, 50% reduction in social contacts and elderly shielding [10]. This covers the range of values considered in this study and demonstrates the potential level of physical distancing together with high-coverage contact tracing to keep the effective reproduction number below one.

We also assume the Negative Binomial dispersion, k, of secondary cases, does not vary with R_S due to different social distancing measures. This relationship is poorly characterised, but it is believed that social distancing may increase k, leading to decreased heterogeneity in number of contacts, potentially making outbreak control harder, although this effect is expected to be at least cancelled out by the reduction in the mean [30]. Furthermore it is also possible that less heterogeneity in contacts may make tracing of individual contacts more feasible, allowing for a higher coverage.

Contact tracing improvements include *secondary contact tracing* as seen in Vietnam, i.e. tracing the contacts of contacts of known cases, to get ahead of the chain of transmission [31]. The use of a digital tracing app across the UK if combined with manual tracing could boost tracing coverage [32] and interactive dashboards are being rolled out across a number of countries to inform modelling efforts and raise public awareness [33]. Backwards contact tracing, whilst highly labour intensive, could also fill vital gaps where transmission links have been missed [34]. As experience in contact tracing

develops, it will also likely be possible to give contacts a prior probability of infection (based on the duration and contact setting for example) and combine this with the test results to give a more accurate measure by which to determine isolation requirements.

Overall, we conclude that contact tracing could bring substantial benefits to controlling and preventing outbreaks, with tracing coverage and speed playing an important role, as well as testing. However, any 'test & trace' strategy must carefully consider the limitation of poor test sensitivity, as well as the additional tracing information obtained from testing asymptomatic individuals. Poorly sensitive tests are inappropriate for ruling out a diagnosis, and infectious individuals immediately halting quarantine following a false negative result could have dangerous implications. In line with previous studies [8], we have demonstrated that contact tracing alone is highly unlikely to prevent large outbreaks unless used in combination with evidence-based physical distancing measures, including restrictions on large gatherings.

Methods

In this extension of a previous COVID-19 branching process model [3], the number of potential secondary cases generated by an index case is drawn from a Negative Binomial. The exposure time for each case, relative to infector onset, is drawn from a shifted Gamma distribution that allows for pre-symptomatic transmission and is left-truncated to ensure secondary case exposure time is after the primary case exposure time. Secondary cases are averted if the primary case is classified as 'quarantined' at the time of infection, assuming within household segregation is possible. The probability of quarantine depends on whether the primary case was traced, any test result, and adherence to self-isolation recommendations (Figure 4). Each simulation was seeded with five infected individuals that are initially undetected by the contact tracing system.

Secondary case distribution

A standard Negative Binomial assumption was used to represent heterogeneity in onward transmission due to factors such as individual contact patterns or infectiousness, with the mean relating to the *effective reproduction number under physical distancing* R_S which takes a value of 1·1, 1·3 or 1·5 with a constant dispersion parameter $k = 0{\cdot}16$, as used in the

original analysis [30, 3]. The estimates of k for SARS-COV-2 are wide-ranging, from $k = 0{\cdot}1(range : 0{\cdot}05 - 0{\cdot}2)$ for pre-lockdown UK [34] to $k = 0{\cdot}25(range : 0{\cdot}13 - 0{\cdot}88)$ for Tianjin, China during lockdown measures [35]. Due to the variation in the literature we have not updated this parameter as 0·16 lies within these ranges and it is not yet possible to derive accurate national estimates of k for post-lockdown scenarios. Here a smaller k represents greater heterogeneity in transmission and results in the majority of index cases leading to no secondary infections, while a small proportion of individuals infect a large number of secondary cases. All parameter estimates and references can be found in Table 1.

Generation interval

The incubation period (time from exposure to symptoms) is assumed to follow a Lognormal distribution with mean 1·43 and standard deviation 0·66 on the log scale [22]. Each new case is then infected at an exposure time drawn from a Gamma-distributed infectivity profile (shape = 17·77, rate = 1·39 day^{-1}, shift = 12·98 days) relative to their infector's symptom onset. If this time is before the infector's exposure then this value is rejected and re-sampled to prevent negative generation intervals. This Gamma distribution has been fitted under these sampling assumptions to serial interval data published by He et al. [23] using the fitdistr package in `R` and our resulting distributions qualitatively match those presented in the original paper (Figure 5). The exposure time is then compared to the isolation times of the infector and cases are averted if the infector is in isolation when the infection event would have happened. For non-averted cases, symptom onset times are then drawn from the Lognormal incubation period distribution and the probability of a case remaining asymptomatic throughout their infected period is fixed at 40% [19, 20].

Contact tracing

New cases are identified either through tracing contacts of known cases or symptomatic individuals self-reporting to the system, which we model as a two-stage process. Firstly, if an individual is symptomatic (i.e. has a fever and/or dry persistent cough) but untraced we assume that a combination of reduced social activity due to illness, and awareness of COVID-19 prevention measures, results in a 90% chance of self-isolation one day after symptom

onset. Secondly, individuals who self-isolate in this way then have a probability of contacting the tracing programme and reporting their symptoms as a potential case, which can be varied in the model.

The assumption of 90% self-isolation relies on high levels of public awareness and draws on evidence from COVID-19 studies in the United States and Israel that suggest 87·3% to 94% of individuals may isolate if they had COVID-19 symptoms [36, 37]. These figures are supported by a US-based 2013 study that reports 72% of respondents would stay at home if they had flu-like symptoms, provisional on access to sick pay [38], without the additional factor of social responsibility introduced by pandemic awareness. However, it is important to note that this figure may still be optimistic, with one study reporting 62·2% of Japanese citizens surveyed went to work within seven days of onset of symptoms [39] and another that 75·1% of individuals living in a UK households with COVID-19 symptoms admit to leaving the house in the last 24 hours [40].

Contact tracing is initiated where an existing case has been identified and isolated. The contacts of that individual are then traced with 40%–100% coverage. If a contact is successfully traced they will always isolate. The time taken to trace and isolate a contact is either one day or drawn from a Uniform distribution of 1–4 days. In the absence of testing, traced contacts are assumed to isolate until non-infectious—approximately 14 days [23]. Any contacts that show symptoms or test positive will have their contacts traced; this continues until no further cases result in transmission chain extinction.

Testing

In simulations that include testing, we assume test sensitivities of 65% or 95% with the lower value representing true sensitivity observed in healthcare settings [14, 15] and the higher value being closer to measurements in controlled conditions [41] and also to demonstrate utility of an alternative testing protocol with higher sensitivity. Due to the nature of the branching process model, only infected individuals are modelled so the impact of test specificity cannot be assessed under these methods, although the implications would be related to programme feasibility rather than efficacy. Current specificity estimates are believed to be reasonably high in comparison [42, 16, 43], with some estimates of close to 100%, but false positive tests could lead to unnecessary negative socioeconomic impact under any scheme requiring quarantine of healthy individuals.

When testing is included in the model, all individuals that either self-report to the contact tracing system (individual A in Figure 4), or are traced contacts (B & D in Figure 4), are tested. From the moment a contact self-reports or is traced, either a zero- or two-day delay is simulated before the test result is returned, chosen to be representative of UK programme targets. If a positive test is returned, the individual's contacts are traced. If a negative test is returned, two different scenarios are explored; either a) immediate departure from quarantine, or b) the individual is asked to complete a precautionary quarantine period (e.g. 7 days from beginning of isolation). Any contacts of a negative-testing case that were successfully identified prior to receiving the test result are still isolated and tested.

No active case detection

A scenario in which there is no active case detection in the community is considered whereby the only detected cases are those who are hospitalised. This is simulated by reducing the case reporting proportion to 6%, reflecting the hospitalisation rate in the UK [44]. Time from symptom onset to hospitalisation is drawn from an Exponential distribution with mean 5·954 days (fitted to data published alongside a modelling study [44]). We then defined the undetected outbreak size as the number of cases that were exposed prior to the first hospitalisation, given an initial seeding of 5 index cases at $t = 0$. We also consider a scenario of 100 index cases to represent a mass super-spreading event, such as the large outbreaks seen in meat-packing plants across Europe [45] or an instance in the SARS outbreak where a flight attendant is thought to have infected more than 100 individuals [46].

Simulation process

Results presented are the combined output of 5,000 simulations for each parameter combination, or scenario, considered. These results are used to derive the probability of a large outbreak given a range of conditions. A *large outbreak* is considered to be 2,000 cases and each simulation is run for a maximum of 300 days. The threshold of 2,000 cases was chosen by running simulations with a maximum of 5,000 cases and noting which of the simulated epidemics that went extinct; 99% of extinction events occurred before reaching 2,000 cases. The model was written in `R` and the code is publicly available in an online GitHub repository (`https://github.com/`

405 `timcdlucas/ringbp`).

Acknowledgments & funding sources

ELD, TCDL, AB, DA, LP, TMP, GM & TDH gratefully acknowledge funding of the NTD Modelling Consortium by the Bill & Melinda Gates Foundation (BMGF) (grant number OPP1184344). The following funding sources are acknowledged as providing funding for the named authors. This research was partly funded by the Bill & Melinda Gates Foundation (NTD Modelling Consortium OPP1184344: GM). This project has received funding from the European Union's Horizon 2020 research and innovation programme - project EpiPose (101003688: PK). Royal Society (RP/EA/180004: PK). Wellcome Trust (210758/Z/18/Z: JH, SA). Views, opinions, assumptions or any other information set out in this article should not be attributed to BMGF or any person connected with them. TC is funded by a Sir Henry Wellcome Fellowship from the Wellcome Trust (reference 215919/Z/19/Z). TMP's PhD is supported by the Engineering & Physical Sciences Research Council, Medical Research Council and University of Warwick (grant number EP/L015374/1). TMP thanks Big Data Institute for hosting him during this work. All funders had no role in the study design, collection, analysis, interpretation of data, writing of the report, or decision to submit the manuscript for publication.

CRediT contribution statement

Conceptualisation: ELD, TCDL, PK, GFM, TDH
Formal Analysis: ELD, TCDL
Funding acquisition: TDH
Investigation: ELD, TCDL, AB, TMP, LP, DA, TC
Methodology: ELD, TCDL
Software: ELD, TCDL, SA, JH
Visualization: ELD, AB
Writing – original draft: ELD, TCDL
Writing – review & editing: All authors

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] South China Morning Post. Coronavirus: China's first confirmed Covid-19 case traced back to November 17. www.scmp.com/news/china/society/article/3074991/coronavirus-chinas-first-confirmed-covid-19-case-traced-back (2020). Accessed 2020-06-02.

[2] The Guardian. Could Covid-19 have reached the UK earlier than thought? www.theguardian.com/world/2020/jun/01/spate-of-possible-uk-coronavirus-cases-from-2019-come-to-light (2020). Accessed 2020-06-02.

[3] Hellewell, J. *et al.* Feasibility of controlling COVID-19 outbreaks by isolation of cases and contacts. The Lancet Global Health (2020).

[4] Keeling, M. J., Hollingsworth, T. D. & Read, J. M. The efficacy of contact tracing for the containment of the 2019 novel coronavirus (COVID-19). medRxiv (2020).

[5] Worldometer. Coronavirus UK summary. www.worldometers.info/coronavirus/country/uk (2020). Accessed: 2020-05-12.

[6] GOV.UK. Coronavirus (COVID-19). www.gov.uk/coronavirus (2020). Accessed 2020-05-12.

[7] Flaxman, S. *et al.* Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. Nature 1–5 (2020).

[8] Giordano, G. *et al.* Modelling the COVID-19 epidemic and implementation of population-wide interventions in Italy. Nature Medicine 1–6 (2020).

[9] Jarvis, C. I. *et al.* Quantifying the impact of physical distance measures on the transmission of COVID-19 in the UK. BMC Medicine **18**, 1–10 (2020).

[10] Davies, N. G. *et al.* Effects of non-pharmaceutical interventions on COVID-19 cases, deaths and demand for hospital services in the UK: a modelling study. The Lancet Public Health (2020).

[11] BBC. Coronavirus: Some return to work as lockdown eases slightly in England. www.bbc.co.uk/news/uk-52642222 (2020). Accessed: 2020-06-02.

[12] UK Department of Health and Social Care. Coronavirus (COVID-19): getting tested. www.gov.uk/guidance/coronavirus-covid-19-getting-tested (2020). Accessed: 2020-05-12.

[13] UK Department of Health and Social Care. Weekly NHS Test and Trace bulletin, England: 23 July to 29 July 2020. Updated 10th August. (2020).

[14] Menni, C. *et al.* Real-time tracking of self-reported symptoms to predict potential COVID-19. Nature Medicine 1–4 (2020).

[15] Kucirka, L. M., Lauer, S. A., Laeyendecker, O., Boon, D. & Lessler, J. Variation in false-negative rate of reverse transcriptase polymerase chain reaction–based SARS-CoV-2 tests by time since exposure. Annals of Internal Medicine (2020).

[16] Zitek, T. The appropriate use of testing for COVID-19. Western Journal of Emergency Medicine **21**, 470 (2020).

[17] Ferretti, L. *et al.* Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing. Science **368** (2020).

[18] Anderson, R. M., Heesterbeek, H., Klinkenberg, D. & Hollingsworth, T. D. How will country-based mitigation measures influence the course of the COVID-19 epidemic? The Lancet **395**, 931–934 (2020).

[19] Mizumoto, K., Kagaya, K., Zarebski, A. & Chowell, G. Estimating the asymptomatic proportion of coronavirus disease 2019 (COVID-19) cases on board the Diamond Princess cruise ship, Yokohama, Japan, 2020. Eurosurveillance **25**, 2000180 (2020).

[20] Lavezzo, E. *et al.* Suppression of a SARS-CoV-2 outbreak in the Italian municipality of Vo. Nature, 1–5 (2020).

[21] Emery, J. C. *et al.* The contribution of asymptomatic SARS-CoV-2 infections to transmission - a model-based analysis of the Diamond Princess outbreak. medRxiv (2020).

[22] Li, Q. *et al.* Early transmission dynamics in Wuhan, China, of novel coronavirus–infected pneumonia. New England Journal of Medicine (2020).

[23] He, X. *et al.* Temporal dynamics in viral shedding and transmissibility of COVID-19. Nature Medicine 1–4 (2020).

[24] Ganyani, T. *et al.* Estimating the generation interval for coronavirus disease (COVID-19) based on symptom onset data, March 2020. Eurosurveillance **25**, 2000257 (2020).

[25] The Guardian. Greece to resume flights from UK on 15 June with strict rules. www.theguardian.com/world/2020/may/31/greece-to-resume-flights-from-uk-on-15-june-with-strict-rules (2020). Accessed: 2020-06-01.

[26] The Straits Times. Coronavirus: Why a double negative test is needed before discharge. www.gov.uk/coronavirus (2020). Accessed: 2020-06-03.

[27] Imai, N. *et al.* Report 3: Transmissibility of 2019-nCoV. Tech. Rep., Imperial College London, UK (2020).

[28] UCL. Just over half of adults strictly sticking to lockdown guidelines as confidence in government falls. www.ucl.ac.uk/news/2020/may/just-over-half-adults-strictly-sticking-lockdown-guidelines-confidence-government-falls (2020). Accessed: 2020-05-31.

[29] Koo, J. R. *et al.* Interventions to mitigate early spread of SARS-CoV-2 in Singapore: a modelling study. The Lancet Infectious Diseases (2020).

[30] Lloyd-Smith, J. O., Schreiber, S. J., Kopp, P. E. & Getz, W. M. Superspreading and the effect of individual variation on disease emergence. Nature **438**, 355–359 (2005).

[31] Le, S. M. Containing the coronavirus (COVID-19): Lessons from Vietnam. https://blogs.worldbank.org/health/containing-coronavirus-covid-19-lessons-vietnam. (2020). Accessed: 2020-06-03.

[32] BBC News. Coronavirus: NHS virus-tracing app downloaded 55,000 times. www.bbc.co.uk/news/uk-england-hampshire-52617236 (2020). Accessed: 2020-06-03.

[33] Dong, E., Du, H. & Gardner, L. An interactive web-based dashboard to track COVID-19 in real time. The Lancet Infectious Diseases **20**, 533–534 (2020).

[34] Endo, A., Abbott, S., Kucharski, A. J., Funk, S. *et al.* Estimating the overdispersion in COVID-19 transmission using outbreak sizes outside China. Wellcome Open Research **5**, 67 (2020).

[35] Zhang, Y., Li, Y., Wang, L., Li, M. & Zhou, X. Evaluating transmission heterogeneity and super-spreading event of COVID-19 in a metropolis of China. International Journal of Environmental Research and Public Health 3705 (2020).

[36] Heffner, J., Vives, M.-L. & FeldmanHall, O. Emotional responses to prosocial messages increase willingness to self-isolate during the COVID-19 pandemic. PsyArXiv (2020).

[37] Bodas, M. & Peleg, K. Self-isolation compliance in the COVID-19 era influenced by compensation: Findings from a recent survey in Israel. Health Affairs **39**, 936–941 (2020).

[38] Kumar, S., Grefenstette, J. J., Galloway, D., Albert, S. M. & Burke, D. S. Policies to reduce influenza in the workplace: impact assessments using an agent-based model. American Journal of Public Health **103**, 1406–1411 (2013).

[39] Machida, M. *et al.* The actual implementation status of self-isolation among Japanese workers during the COVID-19 outbreak. Tropical Medicine and Health **48**, 1–8 (2020).

[40] Smith, L. E. *et al.* Factors associated with adherence to self-isolation and lockdown measures in the UK; a cross-sectional survey. medRxiv (2020).

[41] van Kasteren, P. *et al.* Comparison of seven commercial RT-PCR diagnostic kits for COVID-19. Journal of Clinical Virology **128** (2020).

[42] Nalla, A. K. *et al.* Comparative performance of SARS-CoV-2 detection assays using seven different primer-probe sets and one assay kit. Journal of Clinical Microbiology **58** (2020).

[43] Grassly, N. *et al.* Report 16: Role of testing in COVID-19 control. Imperial College London (2020).

[44] Verity, R. *et al.* Estimates of the severity of coronavirus disease 2019: a model-based analysis. The Lancet Infectious Diseases 669–677, DOI: 10.1016/S1473-3099(20)30243-7 (2020).

[45] Middleton, J., Reintjes, R. & Lopes, H. Meat plants - a new front line in the COVID-19 pandemic (2020).

[46] Gopinath, S., Lichtman, J. S., Bouley, D. M., Elias, J. E. & Monack, D. M. Role of disease-associated tolerance in infectious superspreaders. Proceedings of the National Academy of Sciences **111**, 15780–15785 (2014).

Parameter	**Values**	**Refs**
Number of initial cases	5, 100	range
Effective reproduction number under physical distancing, R_S	1·1, 1·3, 1·5	range
Dispersion of R_S, k	0·16	[30, 3]
Proportion asymptomatic	40%	[19, 20]
Delay: onset to isolation	1 day	
Incubation period (Lognormal)	mean log: 1·43	[22, 23]
Incubation period (Lognormal)	sd log: 0·66	[22, 23]
Infection time (Gamma)	shape: 17·77	[23]
Infection time (Gamma)	rate: 1·39 day^{-1}	[23]
Infection time shift	-13.0 days	[23]
Untraced self-isolation prob.	90%	[36, 37, 38]
Self-reporting probability	0·5–1.0	range
Contact tracing coverage	0–100%	range
Min time to trace contacts	1 day	
Max time to trace contacts	1–4 days	range
Test sensitivity	65%, 95%	[15, 14, 41]
Delay: isolate to test result	0–2 days	range
Isolation duration if -ve test	0–7 days	range
Proportion cases hospitalised	6%	[44]
Onset to hospitalisation (Exp)	mean: 5.95 days	[44]

Table 1: Model parameters values/ranges. Parameters taken from the literature are fixed and for other parameters a range of values are explored.

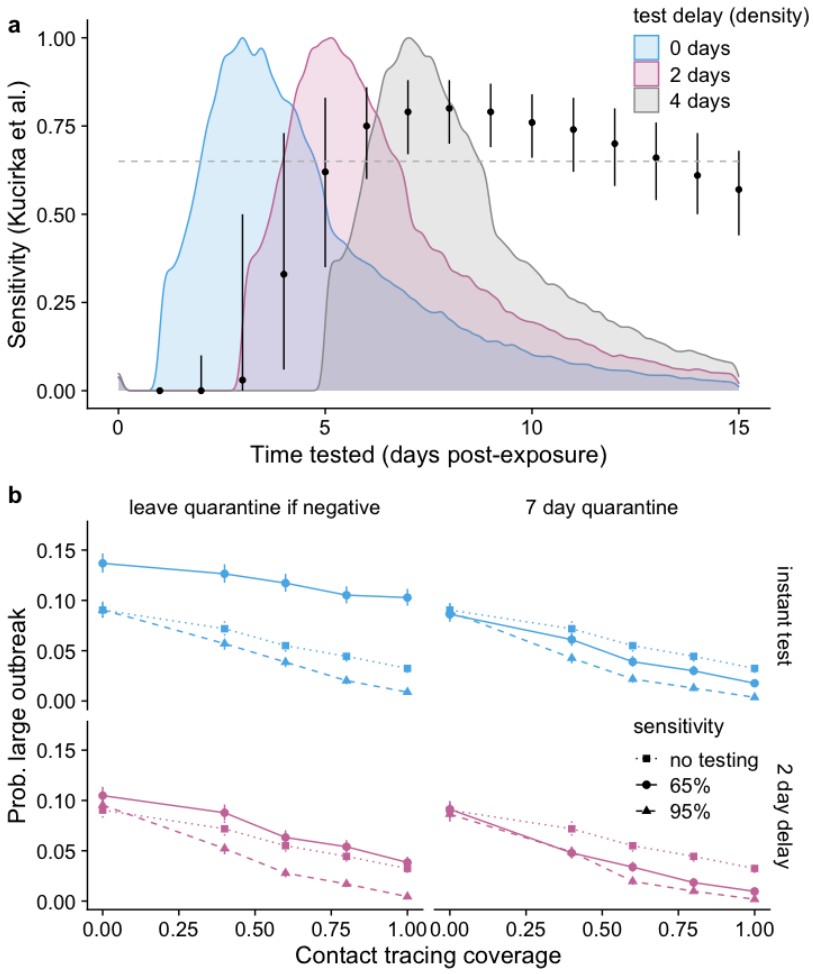


Figure 1: **Test sensitivity and mitigation. a)** Normalised density distributions of time cases are tested in our model, measured in days post-exposure, for an immediate test upon identification (blue), a 2 day delay to testing (pink) and a 4 day delay (grey) for R_S =1·3. Black data points are temporal sensitivity estimates from Kuchirka et al. [15]. Grey dashed line represents 65% sensitivity (as assumed in the model). **b)** Comparing effectiveness of test-and-release of negative-testing cases (left-hand panels) with a minimum seven-day quaratine period (right-hand panels). Assuming 65% sensitivity; 50% self-reporting; 1 day trace delay. Error bars: 95% confidence intervals from simulation output variation.

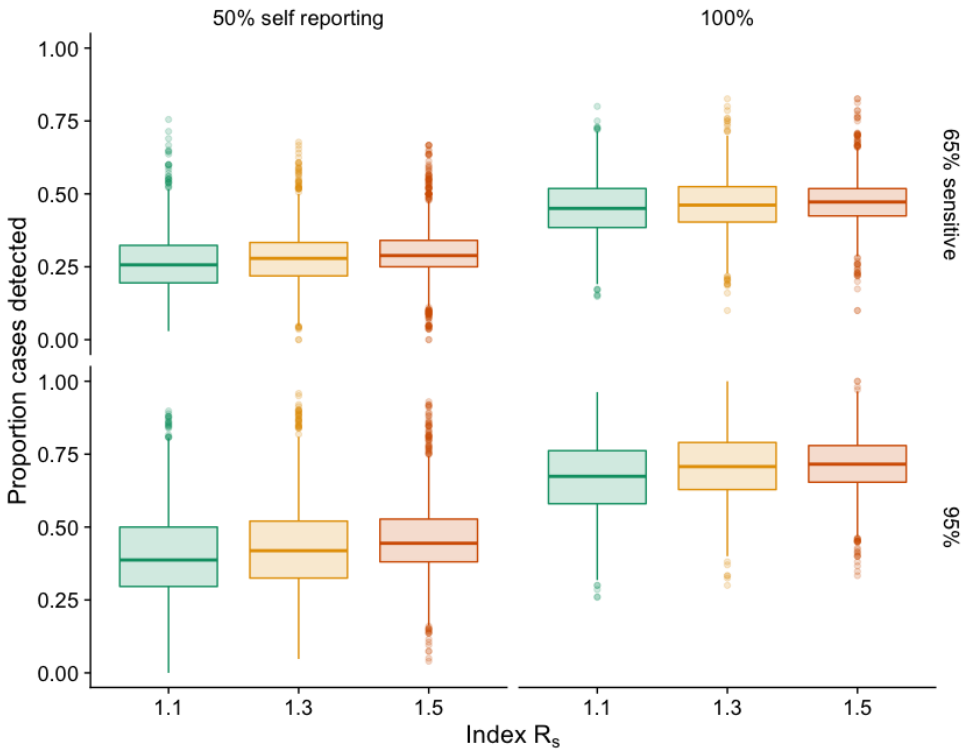


Figure 2: **Case detection.** Proportion of cases detected for varying self-reporting of symptomatic cases (50% and 100%) and diagnostic test sensitivity (65% and 95%). Effective reproduction number under physical distancing: $R_S = 1.1$, 1.3 and 1.5. Box boundaries represent lower (25%), median (50%) and upper (75%) quartiles; whiskers represent the full range of values, excluding outliers, which are marked individually.

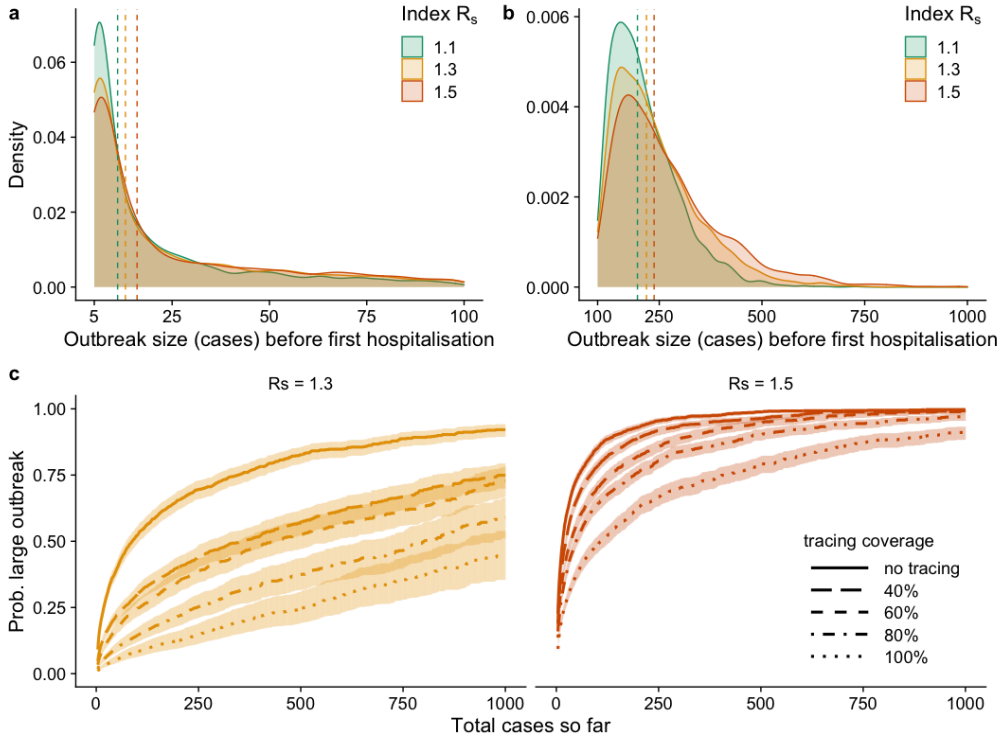


Figure 3: **Super-spreading scenarios.** **a)** Total cases occurring before first hospitalisation in a population with no active tracing or case detection from one super-spreading event (5 new cases). **b)** Same as a) but with 100 new cases. Vertical dashed lines represent median values. **c)** Probability of outbreak by total number of cases so far. Sensitivity = 65%, self-reporting proportion = 50%, individuals testing negative are isolated for a minimum of 7 days, time to test from isolation = 2 days. Error windows: 95% confidence intervals calculated from simulation output variation.

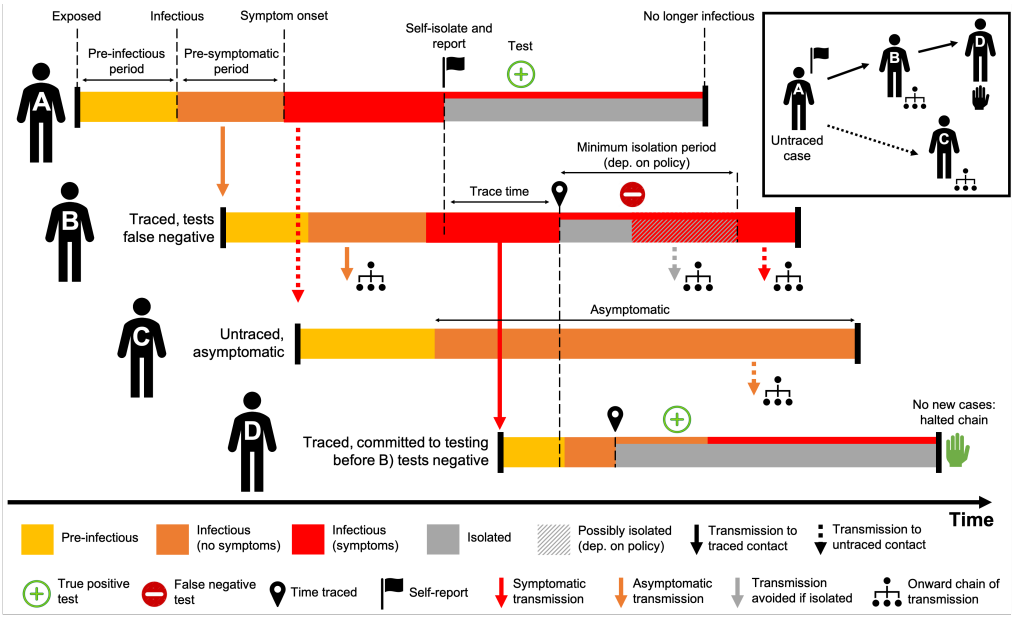


Figure 4: **Contact tracing schematic.** Overview of the contact tracing process implemented in our model. **Person A** isolates and self-reports to the contact tracing programme with some delay after symptom onset, by which time they have infected Persons B and C. When Person A self-reports contact tracing is initiated. They are then tested with positive result and remain isolated for their infectious period. **Person B** was infected by A prior to their symptom onset and is detected by tracing after some delay, after infecting Person D. After isolating they are tested, with a false negative result. This leads to B either a) stopping isolation immediately or b) finishing a minimum 7 day isolation period. Both may allow new onward transmission. **Person C** was infected by A but not traced as a contact. Person C does not develop symptoms but is infectious, leading to missed transmission. **Person D** was traced and tested before the false negative test was returned for Person B. The test for D returns positive, meaning that D remains isolated, halting this chain of transmission.

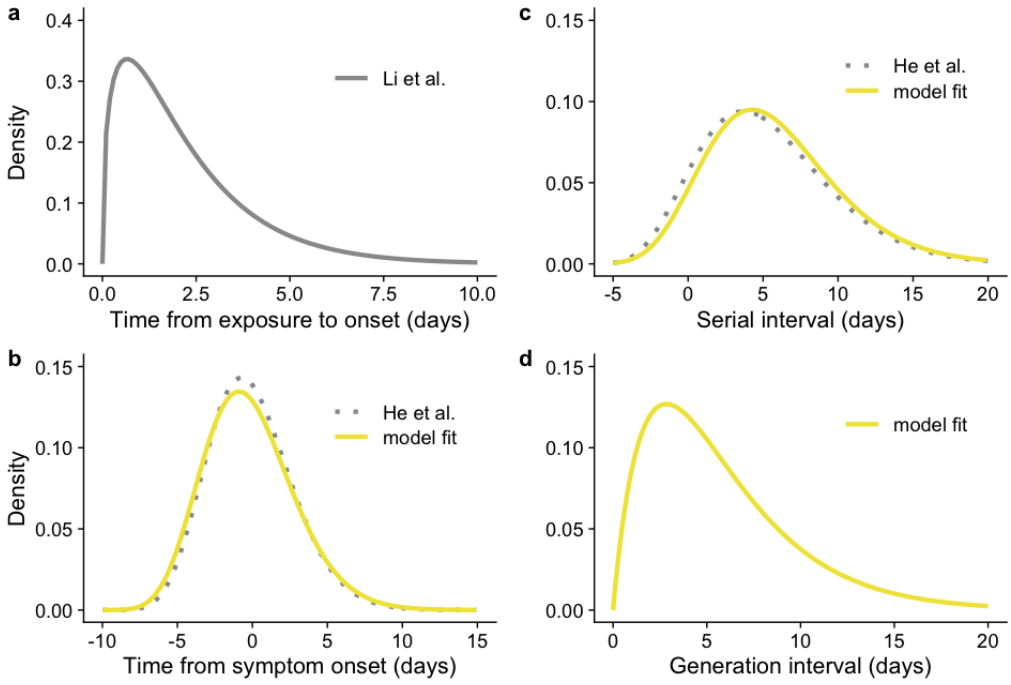
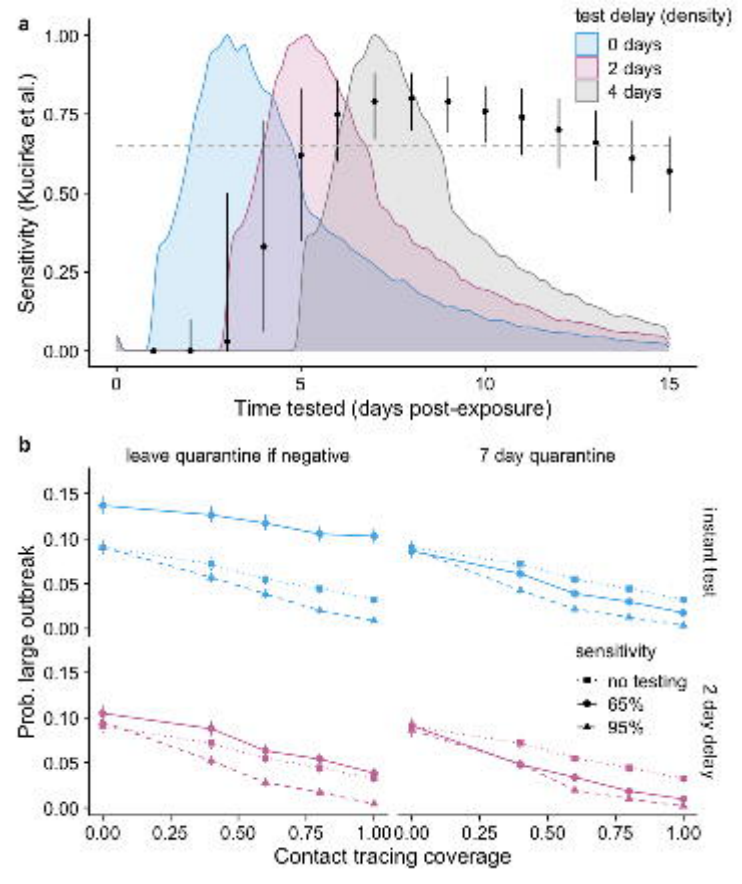
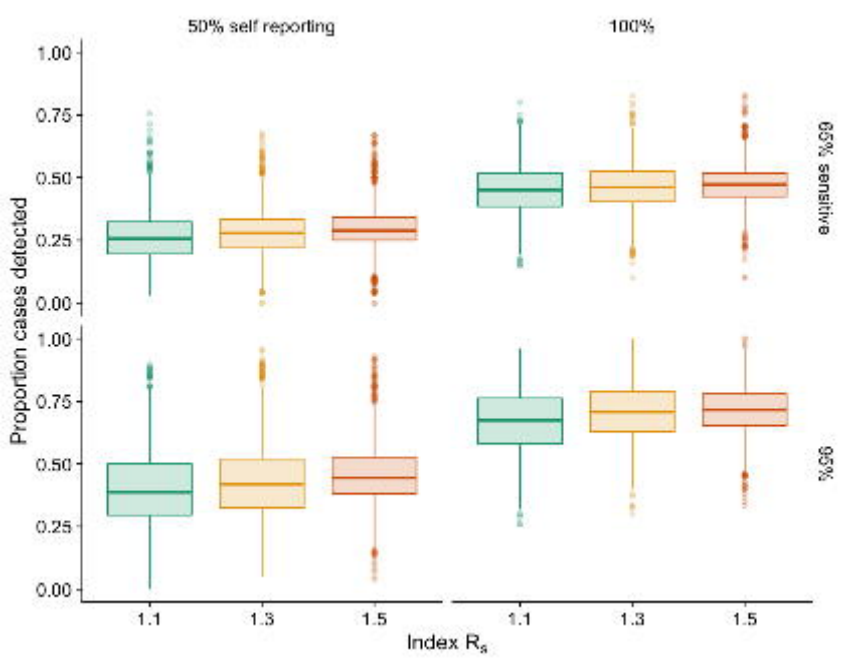
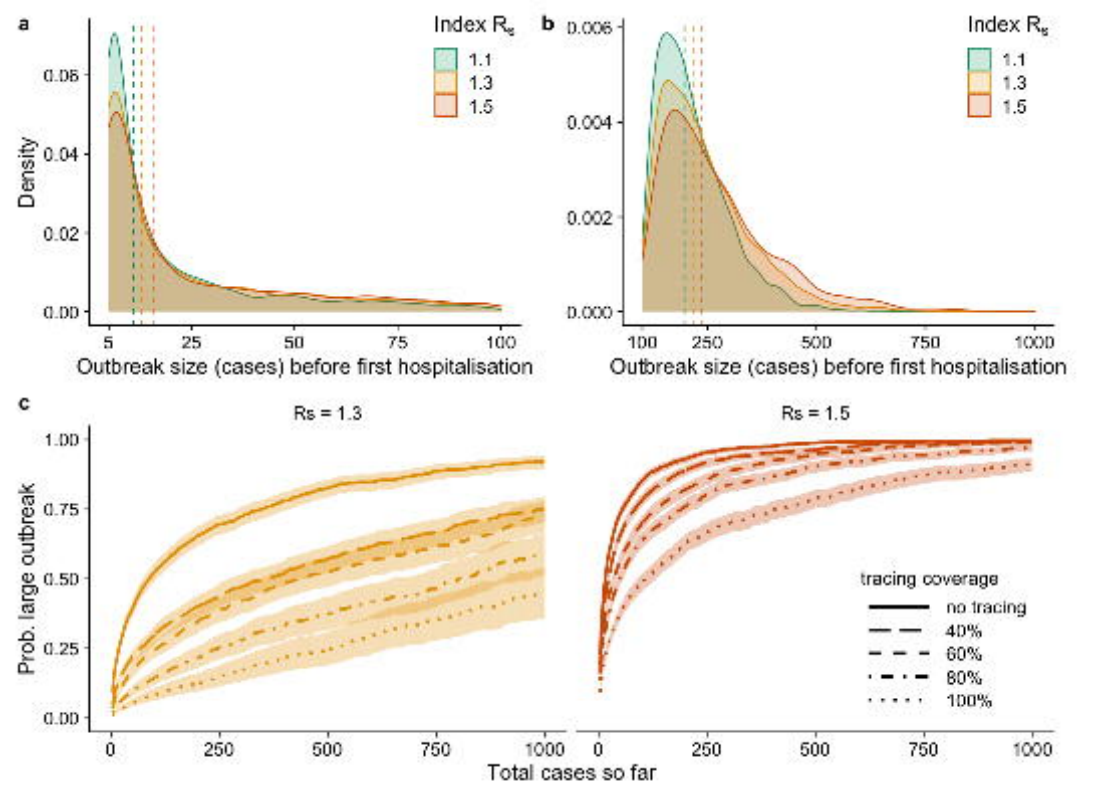
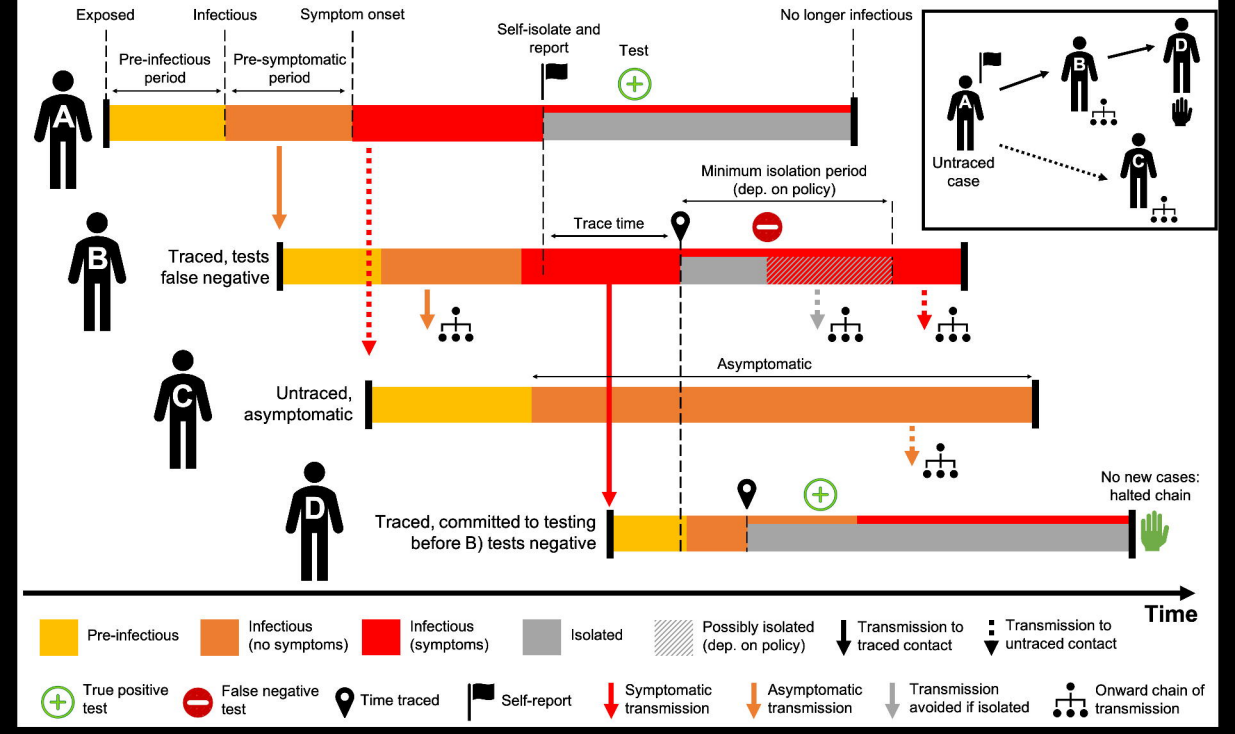


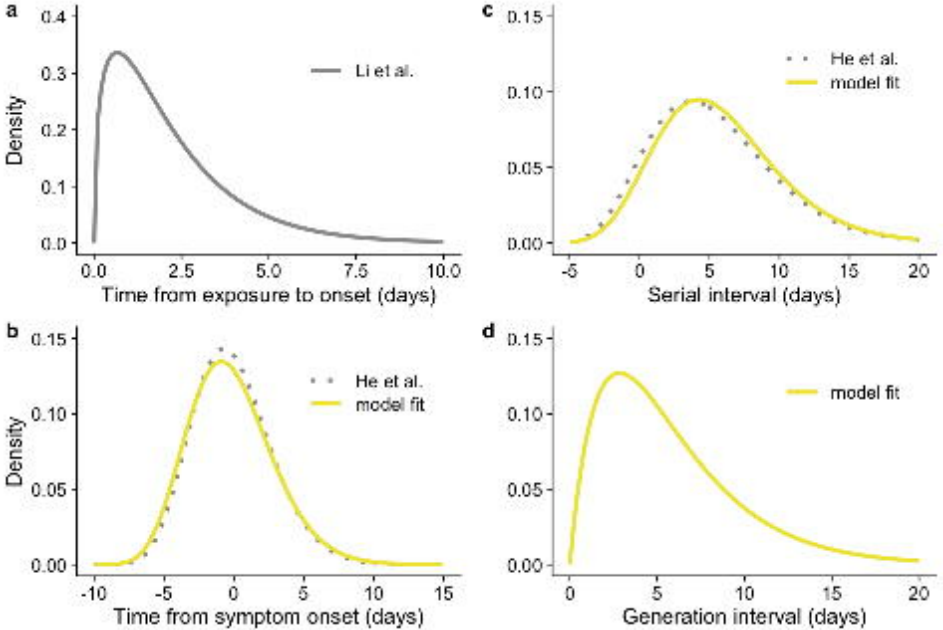
Figure 5: **Parameters' distributions.** Distributions for **a)** incubation period (exposure time to symptom onset) from Li et al. [22]; **b)** transmission profile relative to symptom onset, fitted to data and compared to He et al. [23]; **c)** serial interval, fitted and compared to He et al. [23]; and **d)** generation interval, combined distribution from a) and b) with re-sampling to prevent negative serial intervals, as described in the main text.


a
test delay (density)
0 days
2 days
4 days
Sensitivity (Kucirka et al.)
1.00
0.75
0.50
0.25
0.00
0
5
10
15
Time tested (days post-exposure)
b
leave quarantine if negative
7 day quarantine
instant test
2 day delay
sensitivity
no testing
65%
95%
Prob. large outbreak
0.15
0.10
0.05
0.00
0.00
0.25
0.50
0.75
1.00
Contact tracing coverage


50% self reporting
100%
65% sensitive
95%
Proportion cases detected
1.00
0.75
0.50
0.25
0.00
1.1
1.3
1.5
Index R_s


a
Index R_e
1.1
1.3
1.5
Density
0.06
0.04
0.02
0.00
5
25
50
75
100
Outbreak size (cases) before first hospitalisation
b
0.006
0.004
0.002
0.000
100
250
500
750
1000
Outbreak size (cases) before first hospitalisation
c
Rs = 1.3
Rs = 1.5
Prob. large outbreak
1.00
0.75
0.50
0.25
0.00
tracing coverage
no tracing
40%
60%
80%
100%
0
250
500
750
1000
Total cases so far


Exposed
Infectious
Symptom onset
Self-isolate and report
Test
No longer infectious
Pre-infectious period
Pre-symptomatic period
A
Untraced case
B
C
D
Minimum isolation period (dep. on policy)
Trace time
Traced, tests false negative
Asymptomatic
Untraced, asymptomatic
No new cases: halted chain
Traced, committed to testing before B) tests negative
Time
Pre-infectious
Infectious (no symptoms)
Infectious (symptoms)
Isolated
Possibly isolated (dep. on policy)
Transmission to traced contact
Transmission to untraced contact
True positive test
False negative test
Time traced
Self-report
Symptomatic transmission
Asymptomatic transmission
Transmission avoided if isolated
Onward chain of transmission


a
Density
Li et al.
Time from exposure to onset (days)
b
Density
He et al.
model fit
Time from symptom onset (days)
c
He et al.
model fit
Serial interval (days)
d
model fit
Generation interval (days)