International travel-related control measures to contain the COVID-19 pandemic: a rapid review (Review)

Burns J, Movsisyan A, Stratil JM, Biallas RL, Coenen M, Emmert-Fees KMF, Geffert K, Hoffmann S, Horstick O, Laxy M, Klinger C, Kratzer S, Litwin T, Norris S, Pfadenhauer LM, von Philipsborn P, Sell K, Stadelmaier J, Verboom B, Voss S, Wabnitz K, Rehfuess E

www.cochranelibrary.com

WILEY

TABLE OF CONTENTS

[Rapid Review]

International travel-related control measures to contain the COVID-19 pandemic: a rapid review

Jacob Burns[1,2a], Ani Movsisyan[1,2b], Jan M Stratil[1,2], Renke Lars Biallas[1,2], Michaela Coenen[1,2], Karl MF Emmert-Fees[3], Karin Geffert[1,2], Sabine Hoffmann[1,2], Olaf Horstick[4], Michael Laxy[3,5], Carmen Klinger[1,2], Suzie Kratzer[1,2], Tim Litwin[6], Susan Norris[1,2,7], Lisa M Pfadenhauer[1,2], Peter von Philipsborn[1,2], Kerstin Sell[1,2], Julia Stadelmaier[8], Ben Verboom[1,2], Stephan Voss[1,2], Katharina Wabnitz[1,2], Eva Rehfuess[1,2]

[1]Institute for Medical Information Processing, Biometry and Epidemiology (IBE), Chair of Public Health and Health Services Research, LMU Munich, Munich, Germany. [2]Pettenkofer School of Public Health, Munich, Germany. [3]Institute of Health Economics and Health Care Management, Helmholtz Zentrum München, Munich, Germany. [4]Heidelberg Institute of Global Health, Heidelberg University, Heidelberg, Germany. [5]Department of Sport and Health Sciences, Technical University of Munich, Munich, Germany. [6]Institute for Medical Biometry and Statistics (IMBI), Freiburg Center for Data Analysis and Modeling (FDM), Faculty of Medicine, University of Freiburg, Freiburg, Germany. [7]Oregon Health & Science University, Portland, OR, USA. [8]Institute for Evidence in Medicine, Medical Center, Faculty of Medicine, University of Freiburg, Freiburg, Germany

[a]These authors contributed equally to this work. [b]These authors contributed equally to this work

Contact address: Jacob Burns, burns@ibe.med.uni-muenchen.de.

A B S T R A C T

Background

In late 2019, the first cases of coronavirus disease 2019 (COVID-19) were reported in Wuhan, China, followed by a worldwide spread. Numerous countries have implemented control measures related to international travel, including border closures, travel restrictions, screening at borders, and quarantine of travellers.

Objectives

To assess the effectiveness of international travel-related control measures during the COVID-19 pandemic on infectious disease transmission and screening-related outcomes.

Search methods

We searched MEDLINE, Embase and COVID-19-specific databases, including the Cochrane COVID-19 Study Register and the WHO Global Database on COVID-19 Research to 13 November 2020.

Selection criteria

We considered experimental, quasi-experimental, observational and modelling studies assessing the effects of travel-related control measures affecting human travel across international borders during the COVID-19 pandemic. In the original review, we also considered evidence on severe acute respiratory syndrome (SARS) and Middle East respiratory syndrome (MERS). In this version we decided to focus on COVID-19 evidence only. Primary outcome categories were (i) cases avoided, (ii) cases detected, and (iii) a shift in epidemic development. Secondary outcomes were other infectious disease transmission outcomes, healthcare utilisation, resource requirements and adverse effects if identified in studies assessing at least one primary outcome.

Data collection and analysis

Two review authors independently screened titles and abstracts and subsequently full texts. For studies included in the analysis, one review author extracted data and appraised the study. At least one additional review author checked for correctness of data. To assess the risk of bias and quality of included studies, we used the Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) tool for observational studies concerned with screening, and a bespoke tool for modelling studies. We synthesised findings narratively. One review author assessed the certainty of evidence with GRADE, and several review authors discussed these GRADE judgements.

Main results

Overall, we included 62 unique studies in the analysis; 49 were modelling studies and 13 were observational studies. Studies covered a variety of settings and levels of community transmission.

Most studies compared travel-related control measures against a counterfactual scenario in which the measure was not implemented. However, some modelling studies described additional comparator scenarios, such as different levels of stringency of the measures (including relaxation of restrictions), or a combination of measures.

Concerns with the quality of modelling studies related to potentially inappropriate assumptions about the structure and input parameters, and an inadequate assessment of model uncertainty. Concerns with risk of bias in observational studies related to the selection of travellers and the reference test, and unclear reporting of certain methodological aspects.

Below we outline the results for each intervention category by illustrating the findings from selected outcomes.

Travel restrictions reducing or stopping cross-border travel (31 modelling studies)

The studies assessed cases avoided and shift in epidemic development. We found very low-certainty evidence for a reduction in COVID-19 cases in the community (13 studies) and cases exported or imported (9 studies). Most studies reported positive effects, with effect sizes varying widely; only a few studies showed no effect.

There was very low-certainty evidence that cross-border travel controls can slow the spread of COVID-19. Most studies predicted positive effects, however, results from individual studies varied from a delay of less than one day to a delay of 85 days; very few studies predicted no effect of the measure.

Screening at borders (13 modelling studies; 13 observational studies)

Screening measures covered symptom/exposure-based screening or test-based screening (commonly specifying polymerase chain reaction (PCR) testing), or both, before departure or upon or within a few days of arrival. Studies assessed cases avoided, shift in epidemic development and cases detected. Studies generally predicted or observed some benefit from screening at borders, however these varied widely.

For symptom/exposure-based screening, one modelling study reported that global implementation of screening measures would reduce the number of cases exported per day from another country by 82% (95% confidence interval (CI) 72% to 95%) (moderate-certainty evidence). Four modelling studies predicted delays in epidemic development, although there was wide variation in the results between the studies (very low-certainty evidence). Four modelling studies predicted that the proportion of cases detected would range from 1% to 53% (very low-certainty evidence). Nine observational studies observed the detected proportion to range from 0% to 100% (very low-certainty evidence), although all but one study observed this proportion to be less than 54%.

For test-based screening, one modelling study provided very low-certainty evidence for the number of cases avoided. It reported that testing travellers reduced imported or exported cases as well as secondary cases. Five observational studies observed that the proportion of cases detected varied from 58% to 90% (very low-certainty evidence).

Quarantine (12 modelling studies)

The studies assessed cases avoided, shift in epidemic development and cases detected. All studies suggested some benefit of quarantine, however the magnitude of the effect ranged from small to large across the different outcomes (very low- to low-certainty evidence). Three modelling studies predicted that the reduction in the number of cases in the community ranged from 450 to over 64,000 fewer cases (very low-certainty evidence). The variation in effect was possibly related to the duration of quarantine and compliance.

Quarantine and screening at borders (7 modelling studies; 4 observational studies)

The studies assessed shift in epidemic development and cases detected. Most studies predicted positive effects for the combined measures with varying magnitudes (very low- to low-certainty evidence). Four observational studies observed that the proportion of cases detected for quarantine and screening at borders ranged from 68% to 92% (low-certainty evidence). The variation may depend on how the measures were combined, including the length of the quarantine period and days when the test was conducted in quarantine.

Cochrane
Library

Trusted evidence.
Informed decisions.
Better health.

Cochrane Database of Systematic Reviews

Authors' conclusions

With much of the evidence derived from modelling studies, notably for travel restrictions reducing or stopping cross-border travel and quarantine of travellers, there is a lack of 'real-world' evidence. The certainty of the evidence for most travel-related control measures and outcomes is very low and the true effects are likely to be substantially different from those reported here. Broadly, travel restrictions may limit the spread of disease across national borders. Symptom/exposure-based screening measures at borders on their own are likely not effective; PCR testing at borders as a screening measure likely detects more cases than symptom/exposure-based screening at borders, although if performed only upon arrival this will likely also miss a meaningful proportion of cases. Quarantine, based on a sufficiently long quarantine period and high compliance is likely to largely avoid further transmission from travellers. Combining quarantine with PCR testing at borders will likely improve effectiveness. Many studies suggest that effects depend on factors, such as levels of community transmission, travel volumes and duration, other public health measures in place, and the exact specification and timing of the measure. Future research should be better reported, employ a range of designs beyond modelling and assess potential benefits and harms of the travel-related control measures from a societal perspective.

P L A I N   L A N G U A G E   S U M M A R Y

Can international travel-related control measures contain the spread of the COVID-19 pandemic?

What are international travel-related control measures?

International travel control measures are methods to manage international travel to contain the spread of COVID-19. Measures include:

- closing international borders to stop travellers crossing from one country to another;

- restricting travel to and from certain countries, particularly those with high infection levels;

- screening or testing travellers entering or leaving a country if they have symptoms or have been in contact with an infected person;

- quarantining newly-arrived travellers from another country, that is, requiring travellers to stay at home or in a specific place for a certain time.

What did we want to find out?

We wanted to find out how effective international travel-related control measures are in containing the COVID-19 pandemic.

What we did

We searched for studies on the effects of these measures on the spread of COVID-19. Studies had to report how many cases these measures prevented or detected, or whether they changed the course of the pandemic. The studies could include people of any age, anywhere. They could be of any design including those that used 'real-life' data (observational studies) or hypothetical data from computer-generated simulations (modelling studies).

This is the first update of our review. This update includes only studies on COVID-19, published up to 13 November 2020.

What we found

We found 62 studies. Most (49 studies) were modelling studies; only 13 used real-life data (observational studies). Studies took place across the world and at different times during the pandemic. Levels of COVID-19 within countries varied.

Most studies compared current travel-related control measures with no travel-related controls. However, some modelling studies also compared current measures against possible measures, for example, to see what might happen if controls were more or less relaxed or were combined with other measures.

Main results

Below we summarise the findings of some outcomes.

Travel restrictions reducing or stopping cross-border travel (31 modelling studies)

Most studies showed that travel restrictions reducing or stopping cross-border travel were beneficial, but this beneficial effect ranged from small to large. Additionally, some studies found no effect. Studies also predicted that these restrictions would delay the outbreak, but the delay ranged from one day to 85 days in different studies.

Screening at borders (13 modelling studies and 13 observational studies)

These studies assessed screening at borders, including screening people with symptoms or who had potentially been exposed to COVID-19, or testing people, before or after they travelled.

For screening based on symptoms or potential exposure to COVID-19, modelling studies found that screening reduced imported or exported cases and delayed outbreaks. Modelling studies predicted that 1% to 53% of cases would be detected. Observational studies reported a wide range of cases detected, from 0% to 100%, with the majority of studies reporting less than 54% of cases detected.

For screening based on testing, studies reported that testing travellers reduced imported or exported cases, and cases detected. Observational studies reported that the proportion of cases detected varied from 58% to 90%. This variation might be due to the timing of testing.

Quarantine (12 modelling studies)

All studies suggested that quarantine may be beneficial, but the size of this effect ranged from small to large in the different studies. Modelling studies, for example, predicted that quarantine could lead to between 450 and over 64,000 fewer cases in the community. Differences in effects may depend on how long people were quarantined for and how well they followed the rules.

Quarantine and screening at borders (7 modelling studies and 4 observational studies)

For quarantine and screening at borders, most studies suggested some benefit, however the size of this effect differed between studies. For example, observational studies reported that between 68% and 92% of cases would be detected. Differences in effects may depend on how long people were quarantined for and how often they were tested while in quarantine.

How reliable are these results?

Our confidence in these results is limited. Most studies were based on mathematical predictions (modelling), so we lack real-life evidence. Further, we were not confident that models used correct assumptions, so our confidence in the evidence on travel restrictions and quarantine, in particular, is very low. Some studies were published quickly online as 'preprints'. Preprints do not undergo the normal rigorous checks of published studies, so we are not certain how reliable they are. Also, the studies were very different from each other and their results varied according to the specification of each travel measure (e.g. the type of screening approach), how it was put into practice and enforced, the amount of cross-border travel, levels of community transmission and other types of national measures to control the pandemic.

What this means

Overall, international travel-related control measures may help to limit the spread of COVID-19 across national borders. Restricting cross-border travel can be a helpful measure. Screening travellers only for symptoms at borders is likely to miss many cases; testing may be more effective but may also miss cases if only performed upon arrival. Quarantine that lasts at least 10 days can prevent travellers spreading COVID-19 and may be more effective if combined with another measure such as testing, especially if people follow the rules.

Future research needs to be better reported. More studies should focus on real-life evidence, and should assess potential benefits and risks of travel-related control measures to individuals and society as a whole.

SUMMARY OF FINDINGS

Summary of findings 1.   Travel restrictions reducing or stopping cross-border travel

Disease: COVID-19

Interventions: implementing travel restrictions reducing/stopping cross-border travel; maintaining the measure; early implementation of the measure; implementing a highly stringent measure

Comparators: no measure; relaxation of the measure; late implementation of the measure; implementing a less stringent measure

Outcome	Number of studies	Summary of findings	Certainty of evidence
Outcome category: cases avoided due to measure			
Number or proportion of cases in the community	**13 modelling studies**	Ten out of 13 studies reported reductions in the number or proportion of cases resulting from various travel restrictions. These positive effects ranged from a 1.8% (95% CI -21.9% to 17.5%) reduction to a 97.8% reduction. The remaining three studies reported mixed effects, including a positive effect,no effect or even a negative effect. The variation in the magnitude of effect might be explained by the level of community transmission, implementation of community-based interventions, and the countries restricted by the measure.	**Very low** [a,b,c] $\oplus\circ\circ\circ$
Number or proportion of imported or exported cases	**9 modelling studies**	Eight out of nine studies reported reductions in importations or exportations. These positive effects ranged from a 18% reduction to a 99% reduction. One study reported mixed effects, observing both positive effects and no effect. The variation in the magnitude and direction of effect might be explained by differences in travel volumes, the timing of implementation, the comprehensiveness and severity of the measure implemented.	**Very low** [b,c,d] $\oplus\circ\circ\circ$
Number or proportion of deaths	**3 modelling studies**	All studies showed reductions in deaths. These positive effects ranged from a 4.3% (95% CI -39.1% to 39.1%) reduction to a 98% reduction in deaths. The variation in the magnitude of effect across studies might be explained by differences in the implementation of community-based interventions.	**Very low** [b,c,e] $\oplus\circ\circ\circ$
Risk of importation or exportation	**3 modelling studies**	Two studies reported reductions in the risk of importing and/or exporting cases as a result of travel restrictions; however, no effect estimates were available. The other study reported mixed effects, including an increased risk of importation at some airports, but decreased risk at other airports as a result of lessening travel restrictions. One study suggested that connectedness to the international travel network and the level of community transmission might explain that variation in the effect direction.	**Very low** [c,f,g] $\oplus\circ\circ\circ$
Outcome category: shift in epidemic development			
Probability of eliminating the epidemic	**1 modelling study**	The study reported mixed effects: the probability would be higher (66% probability) for border restrictions followed by strict community measures than for a delayed border closure (55% probability), and the same as early implementation of border restrictions (66% probability).	**Very low** [h,i,j] $\oplus\circ\circ\circ$
Effective reproduction number	**2 modelling studies**	One study reported a beneficial change (i.e. break point) in R_t after the implementation of travel restrictions in European	**Very low** [c,e,i]

		Union countries (mean duration 12.6 days). The other study reported mixed effects, suggesting that complete border closures would lead to a 0.045 reduction in Rt, partial relaxation through the opening of land borders would lead to a 0.177 increase in R_t, while further relaxation allowing for international travel followed by quarantine upon arrival would not lead to a change in R_t.	⊕◯◯◯
Time to outbreak	**6 modelling studies**	Four out of six studies reported reductions in the time to outbreak. These positive effects ranged from a delay of less than one day to 85 days. Two studies reported mixed effects, suggesting both positive effects and no effect. The variation in the direction and magnitude of effect across studies might be explained by differences in the levels of community transmission, the timing of implementation, and the countries restricted by the measure.	**Very low**[b,c,d] ⊕◯◯◯
Risk of outbreak	**2 modelling studies**	One study reported reductions in the risk of an outbreak resulting from travel restrictions with effects ranging from a 1% to a 37% reduction. The other study reported mixed effects, including both a positive effect and no effect. The variation in the magnitude and direction of effect might be explained by differences in the levels of community transmission, the number of cases in the country of departure, the severity of the travel restriction, co-interventions, and the percentage of contacts being traced.	**Very low** [c,i,j] ⊕◯◯◯
Number or proportion of cases at peak	**2 modelling studies**	Both studies reported reductions in the number or proportion of cases at peak. These positive effects ranged from a 0.3% reduction to a 8% reduction. The variation in the magnitude of effect might be explained by differences in the implementation of community-based interventions.	**Low**[k,l] ⊕⊕◯◯
Epidemic growth acceleration	**1 modelling study**	The study reported that international travel controls would lead to a decrease in the growth acceleration of the epidemic progression across 62 countries (−6.05% change, P < 0.0001).	**Low** [h,m] ⊕⊕◯◯
Exportation growth rate	**1 modelling study**	The study reported that both the lockdown of Hubei, resulting in a ban of all travel, as well as travel restrictions imposed on China led to a decrease in the growth rate of cases exported from Hubei and the rest of China, to the rest of the world.	**Low** [h,m] ⊕⊕◯◯

Outcome category: cases detected due to the measure

No contributing study

[a]Downgraded -1 for risk of bias, due to major quality concerns in some studies related to the appropriateness of the model's structural elements, the input parameters, and the adequacy of assessment of the model's uncertainty.

[b]Downgraded -1 for imprecision, due to a wide range of plausible effects.

[c]Downgraded -1 for indirectness, due to no reporting of external validation in some studies and/or concerns with reporting of external validation in others.

[d]Downgraded -1 for risk of bias, due to major quality concerns in some studies related to the appropriateness of the model's structural elements, the input parameters, the adequacy of assessment of the model's uncertainty, and incomplete technical documentation.

[e]Downgraded -1 for risk of bias, due to major quality concerns in some studies related to the appropriateness of the model's structural elements and the adequacy of assessment of the model's uncertainty.

[f]Downgraded -1 for risk of bias, due to major quality concerns in some studies related to the appropriateness of the model's structural elements, the adequacy of assessment of the model's uncertainty and the lack of technical documentation.

[g]Downgraded -1 for imprecision, due to effect estimates being unavailable.

[h]Downgraded -1 for imprecision, due to only one contributing study.

[i]Downgraded -1 for imprecision, due to insufficient data reported to enable assessment of precision.
[j]Downgraded -1 for risk of bias, due to major quality concerns in some studies related to the appropriateness of the model's structural elements and input parameters.
[k]Downgraded -1 for risk of bias, due to major quality concerns in some studies related to the appropriateness of the models's structural elements.
[l]Downgraded -1 for indirectness, due to no reporting of external validation in all included studies.
[m]Downgraded -1 for risk of bias, due to major quality concerns in some studies related to the adequacy of assessment of the model's uncertainty.

Summary of findings 2. Screening at borders

Disease: COVID-19

Interventions: implementing entry and/or exit symptom/exposure-based screening; implementing entry and/or exit test-based screening; implementing a highly stringent screening measure

Comparators: no measure; implementing an alternative measure; implementing a less stringent screening measure

Outcome	Number of studies	Summary of findings	Certainty of evidence
Symptom/exposure-based screening at borders			
Outcome category: cases avoided due to the measure			
Number or proportion of cases exported	1 modelling study	The study reported that putting screening measures in place across the world would reduce the number of cases exported per day from China would be reduced by 82% (95% CI 72% to 95%), under the assumption of only 35.7% of symptomatic individuals being detected.	Moderate[a] ⊕⊕⊕◯
Outcome category: shift in epidemic development			
Time to outbreak	4 modelling studies	All studies reported that entry and/or exit screening alone would delay an outbreak. These positive effects ranged from 2.7-day delay (from 45 days to 47.7 days in reaching 1000 cases) to 0.5-year delay (from 1.7 years (95% CI 0.04 to 6.09) to 2.2 years (95% CI 0.6 to 8.11)). The variation in the magnitude of effect might be explained by differences in the timing of implementation, the number of arriving travellers, the percentage of asymptomatic cases screened, and the sensitivity of screening.	Very low[b,c,d] ⊕◯◯◯
Risk of outbreak	1 modelling study	The study reported that under the assumption of one infected person entering Mauritius per 100 days, entry screening with 100% sensitivity would reduce the probability of an outbreak within 3 months to 10% and screening with 50% sensitivity would reduce the probability to 48%.	Low[a,b] ⊕⊕◯◯
Outcome category: cases detected due to the measure			
Number or proportion of cases detected	4 modelling studies	All studies reported reductions in the number or proportion of cases detected. These positive effects ranged from detecting 0.8% (95% CI 0.2% to 1.6%) of cases to detecting 53% (95% CI 35% to 72%) of cases. The variation in the magnitude of effect might be explained by the time window in which the exposure may have occurred, flight duration, the percentage of asymptomatic cases in the population, the combination of entry and exit screening measures, and the sensitivity of screening.	Very low[b,c,e] ⊕◯◯◯

Proportion of cases detected	9 observational studies	Across studies, the proportion of cases detected by entry and/or exit screening measures ranged from 0 to 100%. For symptom and temperature screening, one study reported that the measure detected 100% of cases; however, all other studies reported substantially lower proportions of cases detected, ranging from 0% to 53%. Across studies, the variation in effects could be due to the specific measure; for example, some symptom/exposure screening procedures may have been more thorough than others.	**Very low[c,f,g]** ⊕○○○
Positive predictive value (PPV)	6 observational studies	The PPV ranged from 0 to 100% in studies assessing symptom/exposure screening. This is likely highly dependent on how exactly symptoms are defined in studies, however this is poorly described in most included studies.	**Very low[c,f,g]** ⊕○○○

Test-based screening at borders

Outcome category: cases avoided due to the measure

Proportion of secondary cases	1 modelling study	The study reported that PCR testing all incoming travellers upon arrival, followed by isolation of test-positives and requiring a negative test at the end of the isolation would lead to a reduction in secondary cases of 88% (95% CI 87% to 89%) for a 7-day isolation period and 92% (95% CI 92% to 93%) for a 14-day isolation period.	**Very low[a,e,h]** ⊕○○○
Proportion of imported cases	1 modelling study	The study reported that PCR testing all incoming travellers upon arrival, followed by isolation of test-positives and requiring a negative test at the end of the isolation would lead to a reduction of 90% of imported cases for a 7-day isolation period and 92% for a 14-day isolation period. Testing all incoming travellers and refusing entry to test-positives would lead to a reduction of 77%.	**Very low [a,e,h]** ⊕○○○

Outcome category: shift in epidemic development

No contributing study.

Outcome category: cases detected due to the measure

Days at risk of transmitting the infection into the community	2 modelling studies	Both studies showed that a single PCR test upon arrival would reduce the days that travellers, upon release, remain at risk of transmitting the infection into the community. These positive effects ranged from 0.1 fewer days to 0.3 fewer days at risk of transmission.	**Low [e,i]** ⊕⊕○○
Proportion of cases detected	5 observational studies	The proportion of cases detected ranged from 58% to 90%. The timing of certain procedures could play a role in the variation of effect, with PCR tests conducted two days after arrival potentially being more effective in detecting cases than those conducted immediately upon arrival.	**Low[c,g]** ⊕○○○
Probability of releasing an infected individual into the community	2 modelling studies	Both studies showed reductions in the probability of releasing an infected individual into the community as a result of PCR testing. These positive effects included a risk ratio of 0.55 (95% CI 0.28 to 0.83) and probabilities of releasing an infected individual ranging from 48% to 53% for scenarios with different risks of transmission while travelling.	**Low[c,e]** ⊕⊕○○

![Cochrane Library logo] Trusted evidence.
Informed decisions.
Better health.

Cochrane Database of Systematic Reviews

[a]Downgraded -1 for imprecision, due to only one contributing study.

[b]Downgraded -1 for risk of bias, due to major quality concerns in some studies related to the appropriateness of the model's structural elements, the input parameters, and the adequacy of assessment of the model's uncertainty.

[c]Downgraded -1 for imprecision, due to a wide range of plausible effects

[d]Downgraded -1 for indirectness, due to no reporting of external validation in some studies and concerns with reporting of external validation in others.

[e]Downgraded -1 for indirectness, due to no reporting of external validation in all included studies.

[f]Downgraded -1 for risk of bias, due to concerns with traveller selection, the reference test, and the flow and timing of procedures.

[g]Downgraded -1 for indirectness, as travellers on evacuation flights and cruise ships comprised most of the studies; these are likely not representative of usual travels.

[h]Downgraded -1 for risk of bias, due to major quality concerns in some studies related to the appropriateness of the model's structural elements and the adequacy of assessment of the model's uncertainty.

[i]Downgraded -1 for risk of bias, due to major quality concerns in some studies related to the adequacy of assessment of the model's uncertainty.

Summary of findings 3.   Quarantine

Disease: COVID-19

Interventions: implementing quarantine; implementing a highly stringent quarantine

Comparators: no measure; implementing an alternative measure (e.g. screening); implementing a less stringent quarantine

Outcome	Number of studies	Summary of findings	Certainty of evidence
Outcome category: cases avoided due to the measure			
Number or proportion of cases in the community	**3 modelling studies**	All studies reported reductions in the number or proportion of cases. These positive effects ranged from 450 fewer cases to 64028 fewer cases during the first wave of the pandemic. The variation in the magnitude of effect might be explained by differences in the population group targeted by the measure.	**Very low**[a,b,c] ⊕◯◯◯
Proportion of imported cases	**1 modelling study**	The study reported that quarantining all incoming travellers would reduce the proportion of imported cases by 55% for a 7-day quarantine period and by 91% for a 14-day quarantine period.	**Very low**[b,d,e,f] ⊕◯◯◯
Number or proportion of cases seeded by imported cases	**3 modelling studies**	All studies reported reductions in the number or proportion of cases seeded by imported cases as a result of quarantine of travellers. These positive effects ranged from a 26% (95% CI 19% to 37%) reduction to a 100% reduction. The variation in the magnitude of effect might be explained by enforcement of the quarantine, age, and the length of the quarantine period.	**Very low** [c,g,h] ⊕◯◯◯
Probability of an imported case not infecting anyone	**1 modelling study**	The study reported that a 14-day quarantine of all international arrivals in New Zealand would lead to a 4% increase in probability in adults and a 14% in the elderly that an imported case would not infect anyone among adults and the elderly. The increase in the probably would be larger when a 14-day government-mandated quarantine is required (31% and 36% among adults and the elderly, respectively).	**Very low** [e,f,i] ⊕◯◯◯
Outcome category: shift in epidemic development			
Time to outbreak	**1 modelling study**	The study reported that increasing the effectiveness of quarantine to 80% and 90% from the base case of 75% effectiveness	**Low**[e,b] ⊕⊕◯◯

Cochrane Library
Trusted evidence.
Informed decisions.
Better health.

Cochrane Database of Systematic Reviews

would delay the peak in active cases and deaths by 3.5 and 5.5 days, respectively.

Outcome category: cases detected due to the measure

Days at risk of transmitting the infection into the community	**2 modelling studies**	Both studies reported reductions in the numbers of days that travellers, upon release, remain at risk of transmitting the infection into the community. These positive effects ranged from 0.1 fewer days to 2.1 fewer days at risk of transmission. The variation in the magnitude of effect might be explained by the length of quarantine.	**Low[f,h]** ⊕⊕◯◯
Proportion of cases detected	**1 modelling study**	The study reported that requiring travellers to quarantine upon arrival in the UK would lead to detecting different proportions of cases, with the magnitude increasing with the number of days in quarantine (7-day quarantine: 51% (95% CI 47% to 56%); 14-day quarantine: 78% (95% CI 74% to 82%)). These proportions are higher than those for screening alone (with either thermal imaging scanners or health checks detecting 0.78% and 1.13% of cases, respectively).	**Very low [a,e,f]** ⊕◯◯◯
Probability of releasing an infected individual into the community	**3 modelling studies**	All studies reported reductions in the risk or probability of releasing an infected individual into the community. These positive effects included a risk ratio ranging from 0.00 (95% CI 0.00 to 0.01) to 0.59 (95% CI 0.28 to 0.85) and probabilities of releasing an infected individual ranging from 0% to 85%. The variation in the magnitude of effect might be explained by the length of the quarantine period and the risk of transmission within quarantine settings.	**Very low[f,h,i]** ⊕◯◯◯

[a]Downgraded -1 for risk of bias, due to major quality concerns in some studies related to the adequacy of assessment of the model's uncertainty and incomplete technical documentation.

[b]Downgraded -1 for imprecision, due to insufficient data reported to enable assessment of precision.

[c]Downgraded -1 for indirectness, due to no reporting of external validation in some studies and concerns with reporting of external validation in others.

[d]Downgraded -1 for risk of bias, due to major quality concerns in some studies related to the appropriateness of the models' structural assumptions and adequacy of assessment of the model's uncertainty.

[e]Downgraded -1 for imprecision, due to only one contributing study.

[f]Downgraded -1 for indirectness, due to no reporting of external validation in included studies.

[g]Downgraded -1 for risk of bias, due to major quality concerns in some studies related to the appropriateness of the models' structural assumptions, the input parameters and the adequacy of assessment of the model's uncertainty.

[h]Downgraded -1 for imprecision, due to a wide range of plausible effects.

[i]Downgraded -1 for risk of bias, due to major quality concerns in some studies related to the adequacy of assessment of the model's uncertainty.

Summary of findings 4. Quarantine and screening at borders

Disease: COVID-19

Interventions: implementing quarantine and screening measures combined

Comparators: implementing a single measure of quarantine or screening

Outcome	Number of studies	Summary of findings	Certainty of evidence

Outcome category: cases avoided due to the measure

No contributing study.

___

Outcome category: shift in epidemic development

___

| Time to outbreak | 1 modelling study | The study reported delays in outbreak resulting from combination of screening and quarantine compared with a single measure. Under the assumption of one flight per day (7.1% of normal travel volume) and 50% sensitivity of screening, the time to outbreak would vary greatly for different combinations of measures ranging from 3.5 years (95% CI 0.09 to 12.9) to 34.1 years (95% CI 0.86 to 126) to outbreak. | Very low [a,b,c] ⊕○○○ |

___

Outcome category: cases detected due to measure

___

Days at risk of transmitting the infection into the community	2 modelling studies	Both studies reported that the combination of quarantine and testing would reduce days that travellers, upon release, remain at risk of transmitting the infection into the community compared with a single measure. These positive effects ranged from 0.01 fewer days to 2.0 fewer days at risk of transmission.	Low [b,c] ⊕⊕○○
Probability of releasing an infected individual into community	3 modelling studies	All studies reported positive effects resulting from a combination of screening and quarantine. These positive effects included a reduction in the probability of releasing an infected individual ranging from 2% to 48%. The variation in the magnitude of effect could be explained by the length of the quarantine period, day(s) on which the test is conducted in quarantine or the risk of transmission within quarantine.	Very low [b,c,d] ⊕○○○
Proportion of cases detected	2 modelling studies	Both studies reported that the combination of quarantine and testing would further increase case detection compared with single measures. These positive effects ranged from 41% to 99% of cases detected. The variation in the magnitude of effect may be explained by the length of the quarantine period with longer quarantine and the duration of travel and stay in the country of departure.	Very low [b,c,e] ⊕○○○
Proportion of cases detected	4 observational studies	All studies reported that the combination of quarantine and testing would further increase case detection compared with single measures. The proportion of cases detected ranged from 68.8% to 90.2%. The type of initial exit and/or entry screening could play a role; while most employed a PCR test upon arrival, one study employed symptom screening. Whether travellers in quarantine were monitored for the development of symptoms, and the intensity of this monitoring may also have been important.	Low [b,f] ⊕⊕○○

___

[a]Downgraded -1 for risk of bias, due to major quality concerns in some studies related to the appropriateness of the model's structural assumptions, the input parameters, and the adequacy of assessment of the model's uncertainty.

[b]Downgraded -1 for imprecision, due to a wide range of plausible effects.

[c]Downgraded -1 for indirectness, due to no reporting of external validation in included studies.

[d]Downgraded -1 for risk of bias, due to major quality concerns in some studies related to the adequacy of assessment of the model's uncertainty.

[e]Downgraded -1 for risk of bias, due to major quality concerns in some studies related to the appropriateness of the model's structural assumptions and the adequacy of assessment of the model's uncertainty.

[f]Downgraded -1 for indirectness, as travellers on evacuation flights comprised most of the studies; these are likely not representative of usual travels.

BACKGROUND

Description of the condition

The first case of the novel coronavirus disease 2019 (COVID-19) was reported in Wuhan, Hubei, China in late 2019. Over the following weeks, the disease spread further in China and several other Asian countries, including Japan, South Korea, and Thailand (WHO 2020a). By mid-March 2020, COVID-19 cases had been reported in over 100 countries across the globe. On 11 March 2020, the World Health Organization (WHO) declared the outbreak to be a global pandemic (WHO 2020b).

COVID-19 is caused by SARS-CoV-2, a virus closely related to those of the coronaviruses which cause severe acute respiratory syndrome (SARS-CoV-1/SARS) and Middle East respiratory syndrome (MERS-CoV/MERS). However, in comparison with these viruses, SARS-CoV-2 has higher transmissibility and lower pathogenicity (Fani 2020). Most people infected with SARS-CoV-2 have mild disease with non-specific symptoms (Wu 2020). The proportion of cases becoming critically ill, with respiratory failure, septic shock, multiple organ failure, or a combination of two or all of these, has been reported as 5% in China (Wu 2020). The length of stay in hospital varies from less than one week to nearly two months; the length of stay in intensive care ranges from one to three weeks (Rees 2020). Among hospitalised patients, mortality from COVID-19 is reported to be 20% (95% confidence interval (CI) 18% to 23%), 23% (95% CI 19% to 27%) and 11% (95% CI 7% to 16%) in the USA, Europe and China, respectively (Dorjee 2020). Although long-term research is still lacking, there is also growing concern over 'long COVID', defined as "signs and symptoms that develop during or following an infection consistent with COVID-19 and which continue for more than four weeks and are not explained by an alternative diagnosis" (NICE 2020). Long COVID is likely to affect 10% or more of those who have tested positive for SARS-CoV-2 (Carfi 2020; Greenhalgh 2020). Even a mild course of COVID-19 may be associated with long-term symptoms, most commonly cough, fever and fatigue, but also shortness of breath and chest pain, headaches and neurocognitive difficulties, and various mental health conditions (Greenhalgh 2020). It is estimated that between 4% and 41% of infected individuals never develop symptoms (Byambasuren 2020). Both presymptomatic and asymptomatic transmission have been described and are likely to play an important role in the dynamics of the pandemic (Furukawa 2020).

A range of non-pharmacological interventions (NPIs) have been put into place by governments to contain and mitigate the spread of COVID-19. Given the lack of a drug to prevent SARS-CoV-2 infection, the current stage of vaccine distribution and provision, and the limited pharmacological interventions to treat COVID-19, NPIs will continue to play a critical role in containing the SARS-CoV-2 pandemic for a significant period of time to come. Travel-related control measures, one important type of NPI, range from the screening of travellers entering or leaving a country to the complete closure of national borders. Starting from February 2020, many countries and regions in the world implemented some type of travel-related control measure, and these continue to be implemented across many countries. As the pandemic continues across the globe, with many countries having experienced a second wave of infection, and others having moved beyond this second wave, it is crucial to understand the effectiveness of these measures, including at what point in an outbreak they should be

implemented and when they can be relaxed. Such knowledge will help to inform decisions on implementation or re-implementation, relaxation or suspension of these measures, as well as potential modifications to them, and will help to guide public health resource allocation. This is in line with the WHO's International Health Regulations 2005, which call to ground public health decision-making in scientific evidence (WHO 2005).

Description of the intervention

Travel-related control measures comprise different interventions, including the complete closure of national borders to entry or exit, or both; travel restrictions reducing or stopping cross-border travel (e.g. denial of entry or exit on the basis of nationality, travel history, health status or other characteristics, suspension of travel via air, land, and sea); symptom/exposure-based screening at borders; test-based screening at borders; and quarantine of travellers. These measures can be implemented for all modes of travel, including air, land, and sea.

Travel-related public health measures have a long tradition as a means of preventing the spread of epidemic diseases. Historic examples include the prevention of the spread of bubonic plague through widespread travel-related quarantine in medieval port towns and other locations (Tognotti 2013). More recently, entry screening at national borders was implemented during the SARS epidemic in 2003, and airport exit screening measures were used in efforts to contain the Ebola epidemic in West Africa and the Democratic Republic of Congo between 2014 and 2016 (Mouchtouri 2019).

In 2019, the WHO developed guidelines on non-pharmacological public health measures for mitigating the risk and impact of epidemic and pandemic influenza. Based on a systematic review of the evidence, internal (i.e. subnational) travel restrictions were among the measures recommended during early stages of extraordinary, localised influenza epidemics. In contrast, entry and exit screening were not recommended due to overall ineffectiveness of the measure, and border closures were not recommended, unless required by national law or in extraordinary circumstances (WHO 2019). However, the transmission characteristics of influenza are different from those of SARS-CoV-2 and these insights are therefore not directly applicable to SARS-CoV-2. More directly relevant, two reviews assessed the effectiveness of travel-related control measures in the context of the SARS-CoV-1, MERS-CoV and other infectious disease epidemics (Errett 2020; Mouchtouri 2019). One review reports that effectiveness was limited, as few infected travellers were identified; however, the review finds secondary potential benefits, such as raising awareness and discouraging sick individuals from travelling (Mouchtouri 2019). The second review examined the impact of travel reductions on the spread of infectious diseases other than influenza, and concluded that these had some success in reducing disease spread across countries, but did not halt transmission. It also emphasised the potentially high social, economic, and political costs of travel bans (Errett 2020). Undertaken in the context of the ongoing SARS-CoV-2 pandemic, a Cochrane Rapid Review examined, among other quarantine measures, the effectiveness of quarantining individuals travelling from countries with a declared outbreak (Nussbaumer-Streit 2020). This review found very low-certainty evidence for a small effect for SARS and a potentially larger effect for COVID-19 (Nussbaumer-Streit 2020). Thus, the evidence regarding the effectiveness of travel-related

Cochrane
Library

Trusted evidence.
Informed decisions.
Better health.

Cochrane Database of Systematic Reviews

control measures to prevent infectious disease spread is mixed and incomplete. Importantly, given the different transmission characteristics of influenza and the likely high rate of asymptomatic transmission for SARS-CoV-2 as compared to SARS-CoV-1 or MERS-CoV, many of the insights gained from these other pathogens are not directly transferable. Consequently, a systematic review of the effectiveness of travel-related control measures drawing on the growing evidence base from the COVID-19 pandemic is warranted.

How the intervention might work

Travel-related control measures limit the mobility of potential human carriers of infection when crossing national (and in principle, subnational) borders. These restrictions can be imposed on travellers arriving or leaving via land, air, or sea and are usually implemented by government agencies. The main idea behind all of these measures is to prevent the introduction of an infectious agent (in the present context, SARS-CoV-2) into a country, to reduce or delay the spread of an infectious disease within a country, or both. The intervention thus seeks to achieve a shift in epidemic development, whether by avoiding the epidemic entirely (i.e. cases do not occur at all), by reducing the peak of the epidemic (i.e. fewer cases occur, or are spread over a longer time period) or by delaying the arrival or peak of the epidemic (i.e. cases occur later).

All travel-related control measures are based on the notion that travellers (all travellers or those from specific regions or with specific characteristics) represent a population at risk of being infected and of spreading the infection. For SARS-CoV-2, the risk of an infected person travelling and being unaware of being infected is compounded by the fact that presymptomatic and asymptomatic transmission are likely to play an important role. The intervention works by:

- stopping travel (i.e. complete border closure);
- limiting the number of at-risk individuals entering or exiting a country (i.e. travel restrictions);
- detecting infected individuals based on symptoms or testing for the virus (i.e. symptom/exposure-based screening; test-based screening); and
- preventing disease transmission until a person has been clearly identified as non-infectious (i.e. quarantine).

In light of the high rates of pre- and asymptomatic transmission, certain travel-related control measures may be more appropriate in the SARS-CoV-2 pandemic than others. For example, quarantine of travellers may prove more effective than entry and exit screening.

In addition to their intended positive impact on infectious disease dynamics, travel-related control measures may also have negative health impacts, notably the well-known side effects of quarantine and isolation on mental health. Moreover, they have far-reaching economic, social, legal, ethical, and political implications (Folayan 2015; Nuttal 2014; Nuzzo 2014).

OBJECTIVES

To assess the effectiveness of international travel-related control measures during the COVID-19 pandemic on infectious disease transmission and screening-related outcomes.

METHODS

In May 2020, the WHO asked the review authors to develop an evidence map that would chart the evidence of various travel-related control measures relevant to containing the COVID-19 pandemic (Movsisyan 2021). This map informed the scope and methodological considerations of a subsequent rapid review requested by the WHO. We first published this rapid review in September 2020 (Burns 2020). Because the body of evidence on COVID-19 is growing very quickly, the WHO requested the present (first) update of that review. The methods for the original rapid review were prespecified in a protocol that was submitted to and reviewed by Cochrane (see Appendix 1). The eligibility criteria were reviewed and agreed upon with WHO. The methods used in this update were largely identical to those employed in the original review; we transparently report below any instances where we have adapted the methods.

To conduct this rapid review, we employed abridged procedures of systematic reviewing at certain stages, according to the Cochrane guidance for rapid reviews (Garritty 2020). Specifically, only one review author conducted data extraction, assessed the risk of bias in epidemiological studies and assessed the quality of modelling studies. One review author checked risk of bias and quality ratings of all studies for consistency and plausibility. At least one additional review author checked for the correctness of all data reported in the data synthesis. Two or more review authors discussed any uncertainties during these stages. To ensure that the abridged procedures did not compromise the methodological rigour of the review, but also to ensure that all stages of the review were conducted consistently and correctly, we assigned these data extraction, risk of bias and quality assessment tasks to experienced Cochrane review authors, and involved researchers with modelling expertise to assist with the data extraction and quality assessment of modelling studies. Furthermore, we piloted the procedures for each stage, conducted regular team meetings, and kept a list of rolling questions that were updated continuously.

Criteria for considering studies for this review

Study designs

In the context of a global pandemic, evidence to inform decisions must be generated rapidly, meaning that methods traditionally used to evaluate the impact of interventions, such as randomised controlled trials (RCTs) or quasi-experimental studies, while possible, may not be considered feasible, appropriate, timely or ethical. Indeed, in this specific context, simulation models developed to make predictions about the (highly uncertain) future often represent the only available evidence to guide decision-making. To ensure that we captured all relevant study types, we considered a broad range of empirical studies of any size that provided a quantitative measure of impact, including experimental and quasi-experimental studies, observational studies, and mathematical modelling studies. Thus, we included the following types of studies:

- Experimental and quasi-experimental studies, such as
  * RCTs
  * Interrupted time series (ITS) studies
  * Controlled before-after (CBA) studies and difference-in-differences (DiD) studies
  * Instrumental variable (IV) studies
  * Regression discontinuity (RD) studies
- Observational studies, such as
  * Cohort studies
  * Case-control studies
- Modelling studies, such as
  * Compartmental models (e.g. SEIR-type models comprising multiple compartments, such as S: susceptible, E: exposed, I: infectious, R: recovered)
  * Bayesian hierarchical models (i.e. models comprising several submodels to integrate observed data as well as uncertainty)
  * Spatial models (i.e. modelling disease transmission spatially)
  * Time-series models (i.e. models that model the temporal nature of disease transmission using time-series techniques)

To avoid the inappropriate exclusion of studies, we considered all studies providing a quantitative measure of impact, regardless of whether they were indicated by any of these labels. We considered studies published in peer-reviewed journals as well as those published on preprint servers. Our rationale for including preprint articles was that in the context of a global pandemic, there may be a scientific as well as moral case for publishing studies at the earliest opportunity to inform the emergency response. We included any studies that had been registered but not yet published (in a peer-reviewed journal or on a preprint server) as 'ongoing' studies.

We excluded the following types of studies and publications:

- Case reports
- Studies that did not provide a quantitative measure of impact (e.g. studies providing a graphical summary of the number of cases over time in relation to the introduction of control measures, qualitative studies)
- Diagnostic studies (e.g. assessing the sensitivity and specificity of different screening tests in general; we did, however, include studies on the use of screening tests at national borders as a travel-related control measure)
- Non-empirical studies (e.g. commentaries, editorials, non-systematic literature reviews not reporting primary empirical data)
- Systematic reviews (although relevant reviews were used for backward citation searches)
- Conference abstracts

Population

We included studies on human populations (without any age restriction) susceptible to SARS-CoV-2/COVID-19. To be eligible, modelling studies had to use modelling parameters for disease transmission specified to reflect SARS-CoV-2/COVID-19. In the original review, we also included studies on SARS-CoV-1/SARS and MERS-CoV/MERS (Burns 2020).

For this update, we excluded studies:

- not targeting human transmission;

- concerned with humans at risk of developing other infectious diseases, characterised by different transmission properties (e.g. SARS-CoV-1/SARS and MERS-CoV/MERS, Ebola and viral meningitis, the transmission modes of which are primarily person-to-person, rather than airborne); and
- addressing humans at risk of developing other infectious diseases, for which travel-related control measures do not play a significant role in containing outbreaks (e.g. influenza).

Interventions

We considered travel-related control measures affecting human travel across national borders. We considered both introduction and implementation, as well as relaxation and de-implementation of the following measures.

- Closure of national borders to entry or exit, or both, which stop cross-border travel
- International travel restrictions or bans, or both, which reduce cross-border travel. These may include the following specific measures.
  * Denial of entry or exit, or both, on the basis of nationality, travel history, health status or other characteristics
  * Full or partial suspension of cross-border travel via any or all of land, air and sea
  * Visa requirement or refusal on the basis of nationality, travel history, health status or other characteristics
- Screening at national borders, involving any of the measures listed below, as well as a follow-up measure, such as testing, self-isolation or refusal of entry, only for those who screen positive
  * Temperature measurement (e.g. thermography)
  * Health questionnaire (e.g. symptoms, travel history, contact history)
  * Physical examination
  * Testing for current or past infection
- Quarantine or isolation of travellers crossing national borders, including voluntary or government-mandated quarantine of travellers for different durations and without any follow-up measures, such as testing at certain days of the quarantine
- Any combination of the above measures

We excluded the following types of interventions.

- Combinations of the above-mentioned travel-related control measures with other measures where studies do not provide effect estimates for the travel-related control measures (e.g. studies providing a combined effect estimate for suspension of cross-border travel and use of mandatory face masks in the general population) Studies in which the effect of travel-related control measures cannot be disentangled from the effect of a broader suite of public health measures cannot usefully inform WHO recommendations on whether countries should or should not consider travel-related control measures to contain the COVID-19 pandemic.
- All interventions not directly related to travel, including a range of containment and mitigation measures (e.g. community-based quarantine, personal protective measures, hygiene measures, bans on mass gatherings and other social-distancing measures).
- All interventions related to movement of animals or goods.

- All interventions concerned with human travel across subnational borders. While subnational measures can potentially inform national travel-related control measures, these measures are not prioritised by the WHO. As shown in the previous evidence map (Movsisyan 2021), they are also often impossible to disentangle from other subnational measures, such as lockdowns, community quarantine or social distancing recommendations.
- Travel warnings or travel advice issued by the WHO or national governments.
- Studies of interventions solely concerned with the accuracy of tests rather than their implementation as part of an entry and/ or exit border control measure.
- Studies of interventions related to international travel but not concerned with cross-border impacts, i.e. interventions to contain transmission within closed populations that only assessed their effect on these closed populations (e.g. on cruise ships, within detention centres). This exclusion criterion was added post hoc.
- Usual practice (e.g. seasonal changes to travel) or events (e.g. school holidays) affecting travel but not representing travel-related control measures.
- Cancellation of events affecting international travel but undertaken as a means to prevent mass gatherings (e.g. Hajj, international sporting events, international trade fairs).

We included studies that assessed travel-related control measures as specified above, targeting populations within one country (e.g. the lockdown of Wuhan, China) if their impact was assessed on the population of other countries (e.g. Australia). We additionally considered relevant restrictions between mainland China and Hong Kong and Taiwan, given the existence of a hard border and the implementation of travel-related control measures analogous to those implemented internationally.

Other considerations

There are two Cochrane Rapid Reviews with overlapping studies. One published review focuses on quarantine measures, including quarantining travellers crossing national borders (Nussbaumer-Streit 2020). The other review is concerned with screening measures, including entry and exit screening at national borders (Viswanathan 2020). In discussions with Cochrane and the WHO, we decided that it would be important for decision-makers to be able to access the evidence on all travel-related control measures in a single review. To address the overlap between the present review and the two separately conducted reviews, we checked our review findings with the findings from those reviews. While we identified a few overlapping studies, these are presented and discussed as part of different bodies of evidence and in relation to different scopes. We did not identify any discrepancies in reporting and interpretation. Along with our previous evidence map on travel-related control measures (Burns 2020), we considered these reviews for backward citation searches.

Comparator(s)

We included a range of possible comparators, such as a counterfactual scenario in which the intervention was not implemented, a complete relaxation of the measure, or a partial relaxation of the measure. Likewise, a scenario of no intervention could have been compared against a counterfactual scenario in which an intervention was implemented or relaxed. A relevant study therefore may compare an observed intervention with a simulated scenario of no intervention, while another study may compare simulated stringent interventions with simulated lax interventions, while yet another study may compare an observed intervention with a simulated intervention implemented at an earlier time.

Outcome(s)

Primary outcomes

We considered studies assessing any of the following infectious disease transmission and screening-related outcomes.

- Cases avoided due to the measure (e.g. number, proportion, rate of cases observed or predicted in the community with and without the intervention).
- Shift in epidemic development due to the intervention (e.g. probability of epidemic, time to/delay in epidemic arrival or peak, size of epidemic peak, change in the effective reproduction number).
- Cases detected due to the measure: we focused on outcomes we felt are most relevant for decision-makers in the current pandemic: the proportion of cases detected among the total number of cases (i.e. sensitivity, case detection rate) and the proportion of cases among those screening positive (i.e. the positive predictive value).

Secondary outcomes

We considered the following secondary outcomes if identified in studies that assessed at least one of the primary outcomes.

- Any other infectious disease transmission outcome (e.g. number of severe cases in the community)
- Healthcare utilisation (e.g. number of cases requiring treatment in the intensive care unit (ICU), time until ICU capacity is reached)
- Resource requirements for implementing the intervention (e.g. costs associated with intervention, additional personnel, number of tests required)
- Any adverse effects (e.g. health, economic and social outcomes)
- User acceptability (e.g. passenger confidence)

We did not assess user acceptability in the original review; following exchanges with the WHO, we added this secondary outcome to the update.

Search methods for identification of studies

The search strategy was structured around two blocks focusing firstly on COVID-19, SARS and MERS, and secondly on travel-related control measures. For the first block, we added search terms related to 'test' to make the strategy more sensitive to capturing studies on testing in this update. We conducted the searches in English but aimed to include studies published in any language. The search strategy was informed by the search strategy used in the evidence map for travel-related control measures (Movsisyan 2021). An experienced information specialist adapted and ran the searches, which were verified by a content expert and reviewed by Cochrane.

Electronic databases

For this update, we ran searches in the following electronic databases.

- Ovid MEDLINE and Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Daily and Versions (1946 to 13 November 2020)
- Ovid Embase (1996 to 13 November 2020)

Other searches

We additionally searched the following COVID-19-specific databases.

- Cochrane COVID-19 Study Register (covid-19.cochrane.org), which contains study references from ClinicalTrials.gov, WHO International Clinical Trials Registry Platform (ICTRP), PubMed, medRxiv and other handsearched articles from publishers' websites.
- WHO 'Global literature on coronavirus disease' database (search.bvsalud.org/global-literature-on-novel-coronavirus-2019-ncov), which primarily contains research (published and/or prepublication) articles indexed in PubMed, Web of Science, Global Index Medicus and Embase. In addition, Lanzhou University (Lanzhou, China) submits citations on a daily basis from the China National Knowledge Infrastructure (CNKI) as well as a number of Chinese journal publishers. Due to high overlap across our sources, we added a filter here to exclude records from MEDLINE and Embase.

In the original review, we also searched the US Center for Disease Control and Prevention (CDC) COVID-19 Research Articles Downloadable Database, but this resource is no longer available. Instead, the contents of this database are now contained in both the Cochrane COVID-19 Study Register and the WHO 'Global literature on coronavirus disease' database.

Finally, we conducted backward citation searches of systematic reviews on travel-related control measures known to us or identified through our searches (see Appendix 2) to identify additional eligible studies. The full search strategy is presented in Appendix 3.

Data collection and analysis

Selection of studies

To harmonise the screening process, we asked all review authors involved with the title and abstract screening to screen an initial set of the same 50 studies, after which we organised a group call to discuss any issues. In the original review, one review author screened all titles and abstracts, while a second review author screened only those excluded by the first review author. For this update, we screened all titles and abstracts in duplicate. The team conducted title and abstract screening using the Rayyan online systematic review software (Ouzzani 2016).

As with the title and abstract screening process, we harmonised the full-text screening process by asking all review authors involved with full-text screening to screen an initial same set of 10 studies (Garritty 2020). The team then discussed any open questions or issues in a group call. Subsequently, two review authors working independently each screened the remaining full-text records in duplicate. The two review authors discussed any discrepancies,

and consulted a third review author or the entire author team where necessary until they reached consensus. We recorded reasons for exclusion for all studies excluded at the full-text screening stage.

Inclusion of non-English language studies

We considered studies published in all languages. Within the review team, we were able to consider studies in Armenian, English, French, German, Italian, Russian and Spanish, and sought help with translation for any other languages, where needed. We screened a small number of studies in other languages at the title and abstract screening stage, including some with an English abstract and some written in German and Spanish, however, we did not identify any relevant studies in any languages other than English.

Excluding eligible studies from the analysis

For this update, we made the post hoc decision to exclude from the analysis several studies meeting the review eligibility criteria. During data extraction and synthesis, we found these studies to be less informative or potentially misleading for decision-making. These studies included: (1) observational screening studies with limited data; (2) observational ecological studies; and (3) modelling studies using overly simplistic or theoretical assumptions and presenting abstract findings.

1. Some observational studies evaluating entry and/or exit screening measures reported only limited data regarding the effectiveness of the measure. These studies report, for example, how many individuals have been screened, how many were screened positively, and how many were COVID-19 cases. However, due to the lack of a reference test, the true number of cases is unknown. As a result, these studies provide information on how many cases were detected, but not on how many cases were missed; thus we feel that these studies are not sufficiently informative for decision-makers. We also excluded such studies from the analysis in the original review.

2. Observational ecological studies examine the aggregated impact of various travel-related control measures across countries, and, in principle, such studies could be of interest to decision-makers. However, the aggregated nature of the data places these ecological studies at even higher risk of bias than other observational studies, making them even less able to deliver causal insights. Moreover, interventions and outcomes, and the associated results, tend to be operationalised in a simplified manner across countries. Consequently, we felt that these studies were at high risk of delivering over-simplified and biased results.

3. All modelling studies providing an assessment of the impact of travel-related control measures make some assumptions to simulate the real-world; these assumptions relate to aspects such as the intervention itself, the travel scenario and/or the regions implementing and being restricted by the intervention. Studies in which most of these aspects use simplistic or conceptual assumptions, however, tend to provide abstract findings that cannot readily be interpreted or applied. We feel that mainly theoretical studies are not sufficiently informative for decision-makers.

Data extraction and management

One review author extracted study characteristics and data from all included main studies using a data extraction form in Microsoft

Excel. All extracted data were checked by a second review author. We piloted the data extraction form, using three studies that represented different intervention types and that met the inclusion criteria. Appendix 4 provides the details on the data extraction categories. For studies excluded from the analysis, we extracted descriptive characteristics relating to the PICO elements, as well as a short narrative description of the results. To do so, we used a simplified version of the data extraction form used for the main studies.

In the review protocol (see Appendix 1), we specified that we would consider searching for data from external sources to enhance our understanding of the design features of the travel-related control measures and the stage of the pandemic at the time these were implemented. However, given the lack of comprehensive reporting and the inconsistency of the information provided across these sources (e.g. discrepancies in how WHO reports described the stage of the pandemic in earlier months), and given that this information was largely not applicable to modelling studies, we decided against using these sources.

Assessment of risk of bias in included studies

One review author rated the risk of bias or the quality of each included study, depending on the type of study, and a second review author checked the judgements. The studies excluded from the review analysis were not further assessed at this stage. The team of review authors involved with assessing risk of bias and quality was largely the same for this update as for the original review. Given that one new review author was involved with this step, at the outset we discussed how to correctly and consistently apply each of the tools to one screening study and two modelling studies before beginning the assessment. These review authors discussed any questions or uncertainties that arose during the process.

Given the broad range of study designs, we applied multiple tools in assessing the risk of bias or quality of included studies, with the same tools applied in the original review and the present update. We had planned to use version 2 of the Cochrane 'Risk of bias' tool for experimental studies (Higgins 2019), and ROBINS-I for quasi-experimental and observational intervention studies (Sterne 2016). However, we did not identify any experimental studies. We identified two synthetic control studies, which are generally considered a type of quasi-experimental study. However, given that ROBINS-I was not developed for this type of quasi-experimental study with more sophisticated statistical methods, we assessed these studies with the quality appraisal tool developed for modelling studies, as described below.

To appropriately assess the risk of bias of observational studies evaluating screening at borders, which are more closely related to diagnostic studies than intervention evaluations, we decided post-protocol to apply the Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) tool (Whiting 2011), as also employed in the Cochrane review on screening measures to control COVID-19 (Viswanathan 2020). This tool comprises four domains: participant selection (i.e. passenger/traveller selection, for our purposes), the index test, the reference standard, and the flow and timing. For each of these domains, using a series of signalling questions, we provided a judgment of 'low', 'unclear' or 'high' risk of bias for each study. Additionally, the tool facilitates a concrete assessment of generalisability through considering how the population, index test

and reference standard compares with the aspects of interest in this review. In line with QUADAS-2 guidance, we considered how best to apply the tool to our specific review question. The tool, including the specifications we applied in making judgements, is outlined in Appendix 5.

As described in the original review (Burns 2020), no validated tool is available for assessing the risk of bias of modelling studies. Following the suggestions by (Egger 2017), we developed a bespoke tool for the assessment of modelling studies and selected criteria from a rapid review of the methodological literature (Philips 2006) and two methodological studies (Caro 2014; Egger 2017). The tool comprises the following domains: (i) model structure, (ii) input data, (iii) validation, (iv) uncertainty and (v) transparency. The individual criteria we applied, in the form of signalling questions, are outlined in Appendix 6. We reported each of the criteria separately, that is, we did not combine multiple criteria into a summary score. This also allows for a distinction between 'fatal flaw indicators', notably inappropriate structural assumptions and input parameters, and other aspects of model quality and credibility, such as internal and external model validation (Caro 2014).

Contacting study authors

We contacted study authors to request additional information where unclear or non-reported aspects precluded the assessment of eligibility or inclusion in the data synthesis.

Data synthesis

Given that observational studies provide a measured estimate of effect whereas modelling studies predict such an effect, we treated these as two separate bodies of evidence in the synthesis (see also 'Assessment of certainty of evidence').

Due to substantial heterogeneity across included studies with regard to the setting, population, intervention and other contextual factors, as well as study methods, and as specified in the protocol (Appendix 1), we decided that data were not sufficiently similar to conduct meta-analyses. We therefore synthesised the findings narratively and in tabular form, stratified by intervention type and outcome. We adhered to the 'Synthesis without meta-analysis' (SWiM) in systematic reviews reporting guideline (Campbell 2020).

Part one of the narrative synthesis comprised four steps in moving from the effects reported at the individual study-level to a summary across studies: (i) we created a study-by-study table describing the effects of interventions, as well as potential effect moderators, as estimated in each included study; (ii) we classified the effect direction for each reported intervention effect, following recent guidance (Hilton Boon 2020); (iii) for each intervention category and primary outcome, we subsequently looked across contributing studies to develop the summary of findings, including a description of the proportion of studies predicting a positive, negative or no effect for the intervention; (iv) we abstracted this summary of findings for each intervention category and primary outcome into a concise narrative summary to present, along with the certainty of the evidence, in the 'Summary of findings' table and the 'Results' section of the review, paying particular attention to sources of heterogeneity (see below).

Part two involved determining the direction of effect for each intervention-outcome pair, which could be a positive effect, no

effect, mixed effect, or negative effect. For systematic reviews of public health interventions, a beneficial effect of any size beyond the null is often considered to be potentially relevant. Additionally, for travel-related control measures, this minimal important difference is highly context-dependent. For example, the role of international travel in importing cases, and the associated role of travel-related control measures in containing the pandemic, will be different in countries where community transmission is not occurring compared to countries where community transmission is widespread. Consistent with this perspective, we did not consider the size of the effect in determining effect direction.

Specifically, we first specified the comparators used in each study (e.g. measure versus no measure or combined measure versus single measure). In determining effect direction, we classified an effect for which a better outcome was observed for the intervention condition than the comparator condition as 'positive', and an effect for which a worse outcome was observed for the intervention condition as 'negative'. Only studies in which the two conditions reported identical effect estimates were classified as 'no effect'. Many studies assessed an intervention in multiple countries or examined a range of scenarios related to a specific intervention (e.g. in the context of high-, moderate- and low-community transmission). Where studies observed consistent effect directions across these conditions, we classified the effect direction as such; where inconsistent effect directions were observed, we classified the effect direction as 'mixed'.

Assessment of heterogeneity and subgroup analyses

In the absence of meta-analyses, and given the substantial heterogeneity of included studies, we did not conduct analyses of subgroups. As part of our narrative synthesis, however, we aimed to identify potential sources of heterogeneity that may have influenced intervention effectiveness. Given methodological differences, as well as differences in interventions, contexts and outcomes, for modelling studies we focused on potential moderating factors (e.g. level of community transmission, stringency of intervention, level of travel after relaxation of intervention) that were assessed within a given study. The methods used to assess these potential moderators differed widely across individual studies, however we only considered data that were assessed and clearly reported as part of a formal analysis. Given that observational studies of entry and exit screening measures were relatively homogeneous, for these studies we examined potential moderators across studies (e.g. type of screening, timing of polymerase chain reaction (PCR) testing).

Assessment of certainty of evidence

We used the GRADE approach to assess the certainty of the primary outcomes. One review author collated the evidence for each primary outcome and suggested initial certainty of evidence ratings. These were then further deliberated in a team of review authors and a joint decision for certainty of evidence ratings was made for each primary outcome.

The certainty of evidence is defined in GRADE as the extent to which one can be confident that the true effect of an intervention lies on one side of a specified threshold, or within a chosen range (Hultcrantz 2017). In the original review, as well as in this update, we considered 'difference from the null' as an important threshold, assuming that even small effect sizes may be relevant for population-level travel-related control measures, as noted above.

The certainty of evidence rating in GRADE yields four possible levels of evidence: high certainty (i.e. the estimated effect lies close to the true effect), moderate certainty (i.e. the estimated effect is probably close to the true effect), low certainty (i.e. the estimated effect might substantially differ from the true effect), and very low certainty (i.e. the estimated effect is probably substantially different from the true effect).

In accordance with our approach to data synthesis, we rated bodies of evidence from observational and modelling studies separately. In GRADE, evidence from RCTs enters the rating as high certainty, as does evidence from observational studies whose risk of bias has been assessed using ROBINS-I (Schünemann 2019). Subsequently, five domains are used to downgrade evidence, including study limitations, inconsistency, indirectness, imprecision and publication bias, and three domains are used to upgrade evidence, including plausible confounding, large estimates of effect, and dose-response relationship. The upgrading applies only when evidence has not been downgraded.

To rate the certainty of evidence from modelling studies, we used the recent guidance developed by the GRADE Working Group (Brozek 2021). As per the guidance, we initially assessed the body of evidence from modelling studies as high certainty and then used the domains described above to assess certainty of model outputs. We then applied the above domains to further downgrade or upgrade certainty of evidence from modelling studies, using tailored interpretations, as specified in the guidance. For example, risk of bias in modelling studies refers to the credibility of the model and its inputs; inconsistency assesses the difference in the results of two or more models; imprecision examines the model point estimate (e.g. predicted event) and the variability of that estimate; indirectness examines model outputs in relation to the prespecified PICO elements of interest; finally, publication bias refers to the likelihood that relevant models have been developed but not made available (Brozek 2021).

To rate the certainty of evidence from observational studies assessing screening at borders in detecting cases (i.e. the proportion of cases detected and the positive predictive value), we used the GRADE guidance for rating the certainty of evidence for diagnostic tests and strategies (Schünemann 2008). In accordance with the guidance, we initially rated the body of evidence from these cross-sectional studies reporting an appropriate reference standard (e.g. a symptom/exposure-based screening followed by PCR testing) as high-certainty evidence. We then applied the five GRADE domains as described above to further downgrade evidence when deemed appropriate.

R E S U L T S

Results of the search

The PRISMA flow diagram (Moher 2009), shown in Figure 1 describes the study selection process. For this update, we screened 3370 new unique records at the title and abstract screening stage (3033 identified through database searches and 337 through backward citation searches of systematic reviews), in addition to the 3036 unique records screened in the first version (6406 records in total). We screened the full texts of 243 new records, in addition to the 385 records that were screened in the first version (628 records in total). Overall, 88 records met the eligibility criteria for this update (comprising 60 new records in addition to the 28 records focusing

on SARS-CoV-2/COVID-19 included in the original review). Reasons for excluding studies at the full-text screening stage are presented in Figure 1. Ninety-three of these studies, the exclusion of which

was decided in rounds of discussion among the review authors, are further described in the Characteristics of excluded studies tables.

Figure 1. Systematic review PRISMA flow diagram
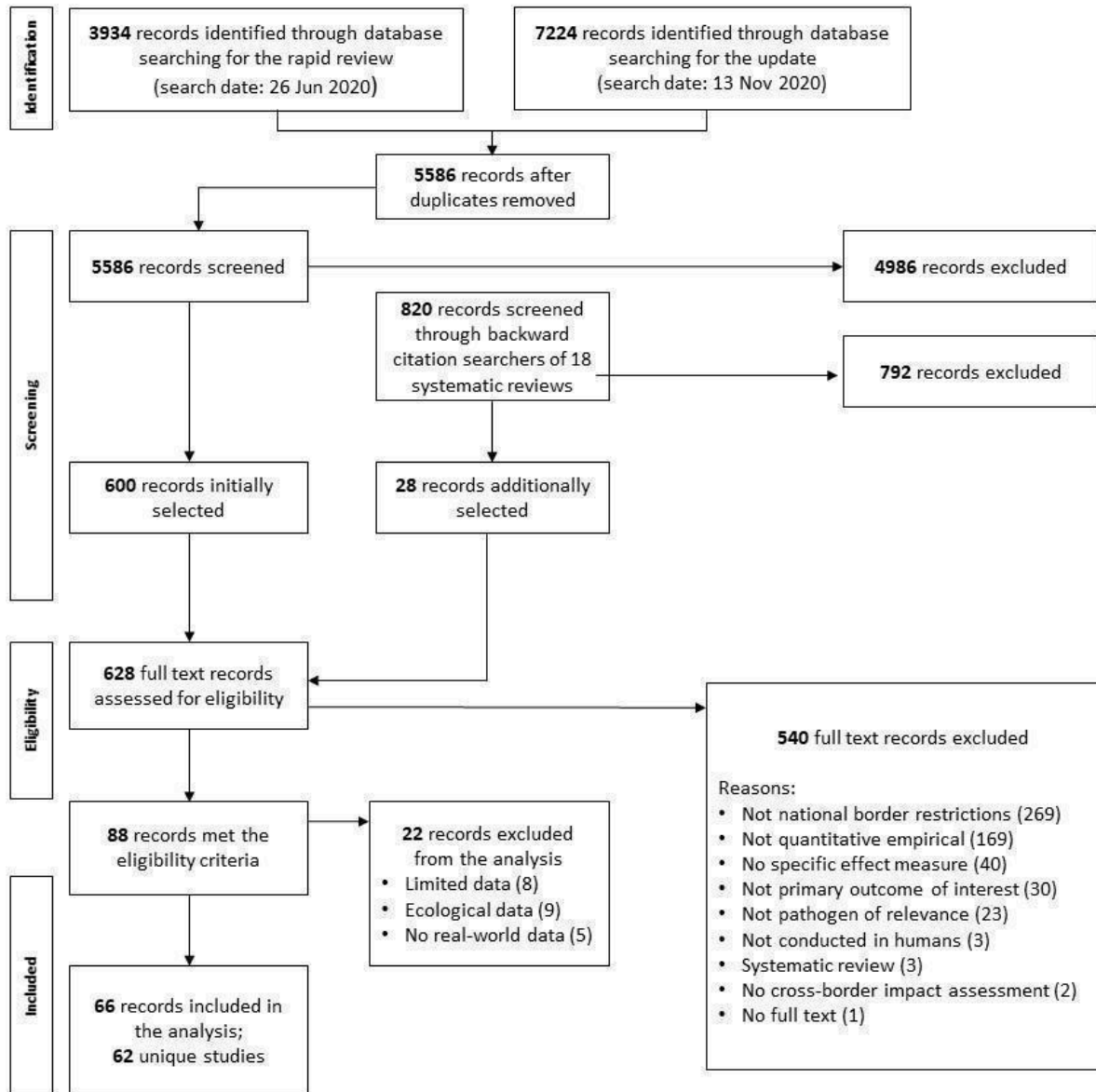
Out of the 88 records, we excluded 22 records from the analysis; thus, these did not contribute effects to the data synthesis or inform conclusions. As described in more detail under 'Methods', these comprised observational screening studies with limited

data (Chang 2020; Expert-Taskforce 2020; Gupta 2020; Hayakawa 2020; Ing 2020; Jernigan 2020; Potdar 2020; Sriwijitalai 2020a), observational ecological studies (Arshed 2020; Chaudhry 2020; Jablonska 2020; Koh 2020; Leffler 2020; Liu 2020a; Ogundokun

2020; Stokes 2020; Teixeira da Silva 2020), and modelling studies for which a 'real-world' effect cannot readily be interpreted (Baba 2020; Chen 2020d; Cacciapaglia 2020a; Cacciapaglia 2020b; Jorritsma 2020). The characteristics of these 22 studies are described in Appendix 7.

We included 66 records in the analysis. These represent 62 unique studies, as four records assessed interventions already addressed by other included records (Arima 2020; Bays 2020; Linka 2020a; Yamahata 2020). The characteristics of each of the 62 studies are described in detail in the 'Characteristics of included studies' and summarised below.

We contacted eight study authors requesting additional information. We did not identify any ongoing studies.

Included studies

The 62 studies included in the analysis are described in the following sections.

Setting

We found studies that evaluated or simulated travel-related control measures in a range of countries across the globe, representing all WHO regions. Countries included Australia (Adekunle 2020; Costantino 2020; Liebig 2020; McLure 2020), Bahrain (Al-Qahtani 2020), Brunei (Wong J 2020), China, including Hong Kong and Macao (Chen J 2020; Chen T 2020; Lio 2020; Pinotti 2020; Kwok 2020; Wells 2020; Wong MC 2020; Yang 2020; Zhang L 2020), France (Lagier 2020), Germany (Hoehl 2020), Greece (Lytras 2020), India (Mandal 2020), Ireland (Grannell 2020), Japan (Arima 2020; Yamahata 2020), Kenya (Kivuti-Bito 2020), Lebanon (Deeb 2020), Malaysia (Shaikh Abdul Karim 2020), Mauritius (Nuckchady 2020), New Zealand (Binny 2020; James 2020; Steyn 2020; Wilson 2020), Saudi Arabia (Al-Tawfiq 2020), Singapore (Chen T 2020; Ng 2020), South Korea (Boldog 2020; Kim 2020; Ryu 2020), Switzerland (Sruthi 2020), Taiwan (Chen Y-H 2020), Thailand (Boldog 2020), UK (Clifford 2020b; Taylor 2020), and the USA (Boldog 2020; Davis 2020; Nowrasteh 2020; Odendaal 2020). Ten studies assessed measures implemented across multiple countries using a cross-country comparison (Anderson 2020; Chinazzi 2020; Kang 2020; Linka 2020a; Nakamura 2020; Russell TW 2020; Shi 2020; Utsunomiya 2020; Zhang C 2020; Zhong 2020), while eight modelling studies did not refer to a specific country or setting (Anzai 2020; Ashcroft 2020; Bays 2020; Clifford 2020a; Dickens 2020; Gostic 2020; Quilty 2020; Russell WA 2020). Most studies specified a travel-related control measure that restricted travel from China (Adekunle 2020; Anzai 2020; Boldog 2020; Chen J 2020; Chinazzi 2020; Costantino 2020; Davis 2020; Hoehl 2020; Kang 2020; Lagier 2020; Liebig 2020; Lio 2020; Mandal 2020; McLure 2020; Ng 2020; Nowrasteh 2020; Odendaal 2020; Pinotti 2020; Ryu 2020; Shaikh Abdul Karim 2020; Shi 2020; Kwok 2020; Wells 2020). Other regions restricted by the travel measures assessed in the studies were Australia (Wilson 2020), Bahrain (Al-Tawfiq 2020), Canada (Al-Tawfiq 2020), Dubai (Al-Tawfiq 2020), Egypt (Al-Tawfiq 2020), Indonesia (Liebig 2020), Iran (Adekunle 2020; Kim 2020; Liebig 2020; Shaikh Abdul Karim 2020), Ireland (Grannell 2020), Italy (Adekunle 2020; Al-Tawfiq 2020; Liebig 2020; Shaikh Abdul Karim 2020; Wong MC 2020), Japan (Wong MC 2020), Oman (Al-Tawfiq 2020), Singapore (Chen J 2020), South Korea (Adekunle 2020; Liebig 2020), Spain (Al-Tawfiq 2020; Lytras 2020), the UK (Al-Tawfiq 2020; Lytras 2020), and the USA (Al-Tawfiq 2020; Linka 2020b).

Population

Sixty-two studies assessed the impact of travel-related control measures in relation to COVID-19 (Adekunle 2020; Al-Qahtani 2020; Al-Tawfiq 2020; Anderson 2020; Anzai 2020; Arima 2020; Ashcroft 2020; Banholzer 2020; Bays 2020; Binny 2020; Boldog 2020; Chen J 2020; Chen T 2020; Chen Y-H 2020; Chinazzi 2020; Clifford 2020a; Clifford 2020b; Costantino 2020; Davis 2020; Deeb 2020; Dickens 2020; Gostic 2020; Grannell 2020; Hoehl 2020; James 2020; Kang 2020; Kim 2020; Lagier 2020; Liebig 2020; Linka 2020a; Linka 2020b; Lio 2020; Kivuti-Bito 2020; Lytras 2020; Mandal 2020; McLure 2020; Nakamura 2020; Ng 2020; Nowrasteh 2020; Nuckchady 2020; Odendaal 2020; Pinotti 2020; Quilty 2020; Russell TW 2020; Russell WA 2020; Ryu 2020; Shaikh Abdul Karim 2020; Shi 2020; Sruthi 2020; Steyn 2020; Taylor 2020; Utsunomiya 2020; Kwok 2020; Wells 2020; Wilson 2020; Wong J 2020; Wong MC 2020; Yamahata 2020; Yang 2020; Zhang C 2020; Zhang L 2020; Zhong 2020).

Intervention and comparisons

Included studies referred to a range of travel-related control measures, which we classified into four categories.

1. **Travel restrictions reducing or stopping cross-border travel:** studies in this intervention category used models to simulate COVID-19 outbreak scenarios (Adekunle 2020; Anderson 2020; Anzai 2020; Banholzer 2020; Binny 2020; Boldog 2020; Chen T 2020; Chinazzi 2020; Costantino 2020; Davis 2020; Deeb 2020; Grannell 2020; Kang 2020; Liebig 2020; Linka 2020a; Linka 2020b; McLure 2020; Nakamura 2020; Nowrasteh 2020; Odendaal 2020; Pinotti 2020; Russell TW 2020; Shi 2020; Sruthi 2020; Utsunomiya 2020; Kwok 2020; Wells 2020; Yang 2020; Zhang C 2020; Zhang L 2020; Zhong 2020). The control measures were often simulated as different levels of reduction in travel volume (e.g. 25% and 75% (Adekunle 2020; Anderson 2020; Anzai 2020; Boldog 2020; Chinazzi 2020; Linka 2020a)). While in practice this may imply a border closure or restriction of travel to varying degrees, such a differentiation would be arbitrary based on the methods used in the studies to simulate these measures. We therefore report these in a combined intervention category.

2. **Screening at borders:** studies in this intervention category comprised observational studies and modelling studies reporting data on symptom/exposure-based screening at borders (e.g. presence of cough and/or fever and/or risk factors (Al-Qahtani 2020; Bays 2020; Clifford 2020b)) and/or test-based screening at borders (e.g. PCR testing (Clifford 2020b; Taylor 2020)) (Al-Qahtani 2020; Al-Tawfiq 2020; Arima 2020; Bays 2020; Chen J 2020; Clifford 2020a; Clifford 2020b; Dickens 2020; Gostic 2020; Hoehl 2020; Lagier 2020; Lio 2020; Kim 2020; Lytras 2020; Mandal 2020; Ng 2020; Nuckchady 2020; Quilty 2020; Russell WA 2020; Shaikh Abdul Karim 2020; Steyn 2020; Taylor 2020; Wells 2020; Wilson 2020; Wong J 2020; Yamahata 2020). This intervention category included screening with a follow-up measure, such as testing, self-isolation or refusal of entry, only for those who screened positive. While a few studies explicitly highlighted the presence of this follow-up measure, many did not. In some of the observational studies, screening is followed by a 14-day quarantine, but in intervention category 2 we treat this quarantine period as a way to identify 'true' cases, rather than as an intervention in its own right; relevant studies are also included in intervention category 4.

3. **Quarantine:** modelling studies in this intervention category assessed voluntary or government-mandated quarantine of

Cochrane
Library

Trusted evidence.
Informed decisions.
Better health.

Cochrane Database of Systematic Reviews

travellers of different duration without any accompanying or follow-up measures, such as symptom/exposure-based screening or testing upon arrival or at certain days of the quarantine (Ashcroft 2020; Chen T 2020; Chen Y-H 2020; Clifford 2020b; Dickens 2020; James 2020; Kivuti-Bito 2020; Russell WA 2020; Ryu 2020; Steyn 2020; Taylor 2020; Wong MC 2020).

4. **Quarantine and screening at borders:** the modelling and observational studies in this intervention category reported data on the combination of quarantine of travellers crossing national borders and screening at borders and/or at different days during quarantine (e.g. day 3, 5, 7, and 14) (Al-Qahtani 2020; Arima 2020; Ashcroft 2020; Bays 2020; Chen J 2020; Clifford 2020b; Russell WA 2020; Shaikh Abdul Karim 2020; Steyn 2020; Taylor 2020; Wilson 2020). In the observational studies, after the combined 14-day quarantine and applied screening measures, there is a final PCR testing before release; in intervention category 4 we treat this final PCR testing as a way to identify 'true' cases, rather than as an intervention in its own right.

Some included studies were inconsistent and sometimes ambiguous in how they labelled and described travel-related control measures. The terms "screening" and "testing", for example, were used inconsistently and often interchangeably without further specification of the procedures. In this review, we use the term "screening" more broadly to refer to any procedure to assess an individual for a potential disease, including an assessment of symptom, exposure and/or testing. Where data allows, we have differentiated between entry and/or exit symptom/exposure-based screening alone (i.e. screening for symptoms, such as fever or cough and/or screening for risk factors or when "screening" was used without further specification of the procedures) and entry and/or exit test-based screening. With regard to testing, most studies specified reverse transcription PCR testing (RT-PCR testing – also referred to as quantitative PCR and a method to measure the amount of RNA) or simply PCR testing, while a few studies did not specify the testing procedure at all. In this review, we use the term "PCR testing" for consistency. Similarly, studies were inconsistent in the use of the terms "quarantine" and "(self-)isolation" and often used them interchangeably. In this review, we therefore use the term "quarantine" to refer to the separation of travellers at risk of developing the disease, and "(self-)isolation" to refer to the separation of confirmed cases (HHS 2020).

For this review, we identified the following intervention-comparator pairs.

1. Travel-related control measure (intervention) versus no travel-related control measure (comparator)
2. Maintaining travel-related control measure (intervention) versus relaxing travel-related control measure (comparator)
3. More stringent travel-related control measure (intervention) versus less stringent travel-related control measure (comparator)
4. Earlier travel-related control measure (intervention) versus later travel-related control measure (comparator)
5. Combined travel-related control measure (intervention) versus single travel-related control measure (control)

Since, in most modelling studies, the comparison to the travel-related control measure was a scenario in which the measure was not implemented, we have used one 'Summary of findings' table per intervention category to describe the evidence and have not split the evidence based on different comparators used. Meanwhile, we have developed a separate 'Summary of findings' table for the evidence from the modelling studies comparing combined measures with a single measure.

In the observational studies concerned with screening and travel-related quarantine, the comparison was the counterfactual of not implementing the measure.

Outcomes

Primary outcomes

We included studies assessing three broad categories of primary outcomes.

1. Cases avoided due to the measure
2. Shift in epidemic development
3. Cases detected due to the measure

For category 1, we identified specific outcomes related to the number or proportion of cases in the community (Anderson 2020; Banholzer 2020; Binny 2020; Chen T 2020; Chen Y-H 2020; Costantino 2020; Deeb 2020; Kang 2020; Linka 2020a; Nowrasteh 2020; Kwok 2020; Wong MC 2020; Yang 2020; Zhang C 2020; Zhong 2020), the number or proportion of imported or exported cases (Adekunle 2020; Anzai 2020; Chen T 2020; Chinazzi 2020; Costantino 2020; Dickens 2020; Liebig 2020; McLure 2020; Russell TW 2020; Wells 2020), the number or proportion of cases seeded by imported cases (Dickens 2020; James 2020; Ryu 2020), the probability of an imported case not infecting anyone (James 2020), the number or proportion of deaths (Binny 2020; Costantino 2020; Kwok 2020), the risk of importation or exportation (Nakamura 2020; Shi 2020; Zhang L 2020), and the proportion of secondary cases (Dickens 2020).

For category 2, we identified outcomes related to the probability of eliminating the epidemic (Binny 2020), the effective reproduction number (Linka 2020a; Sruthi 2020), the time to outbreak (Anzai 2020; Clifford 2020a; Davis 2020; Grannell 2020; Linka 2020b; Kivuti-Bito 2020; Mandal 2020; Nuckchady 2020; Odendaal 2020; Wilson 2020; Zhong 2020), the risk of an outbreak (Anzai 2020; Boldog 2020; Nuckchady 2020), the number or proportion of cases at peak (Binny 2020; Grannell 2020), the epidemic growth acceleration (Utsunomiya 2020), and the exportation growth rate (Pinotti 2020).

Finally, for category 3, we identified outcomes related to days at risk of transmission (Clifford 2020b; Russell WA 2020), the number or proportion of cases detected (Al-Qahtani 2020; Al-Tawfiq 2020; Arima 2020; Bays 2020; Chen J 2020; Gostic 2020; Hoehl 2020; Kim 2020; Lytras 2020; Ng 2020; Quilty 2020; Shaikh Abdul Karim 2020; Taylor 2020; Wilson 2020; Wong J 2020; Yamahata 2020), the positive predictive value (PPV) (Arima 2020; Chen J 2020; Hoehl 2020; Kim 2020; Lytras 2020; Ng 2020; Yamahata 2020), and the probability of releasing an infected individual into the community (Ashcroft 2020; Clifford 2020b; Steyn 2020).

It should be noted that studies were also inconsistent in how they described the specific outcomes. In our classification of the specific outcomes within the broader outcome categories, we have used general terms to enable consistent reporting. For example, we use "outbreak" as a broad term to describe the outcomes labelled in specific studies as "occurrence major epidemic" (Anzai 2020),

![Cochrane Library logo] Trusted evidence.
Informed decisions.
Better health.

Cochrane Database of Systematic Reviews

"beginning of community transmission" (Davis 2020), or "epidemic arrival" (Zhong 2020).

Secondary outcomes

We identified four studies reporting on secondary outcomes related to infectious disease transmission and healthcare utilisation (Ashcroft 2020; Chen Y-H 2020; Steyn 2020; Kwok 2020). For the first category, one study reported on the probability of cases seeded by infected front-line workers at quarantine facilities (Steyn 2020), and two studies reported on the number of people quarantined (Ashcroft 2020; Chen Y-H 2020), with one of the studies reporting a metric defined as "utility of quarantine" and measured as a ratio between the amount of overall transmission prevented and the number of person days spent in quarantine (Ashcroft 2020). For the second category of outcomes, we identified one study reporting on the date on which hospital capacity is reached (Kwok 2020).

Study designs

We identified 49 modelling studies across the four intervention categories (Adekunle 2020; Anderson 2020; Anzai 2020; Ashcroft 2020; Banholzer 2020; Bays 2020; Binny 2020; Boldog 2020; Chen T 2020; Chen Y-H 2020; Chinazzi 2020; Clifford 2020a; Clifford 2020b; Costantino 2020; Davis 2020; Deeb 2020; Dickens 2020; Gostic 2020; Grannell 2020; James 2020; Kang 2020; Liebig 2020; Linka 2020a; Linka 2020b; Kivuti-Bito 2020; Mandal 2020; McLure 2020; Nakamura 2020; Nowrasteh 2020; Nuckchady 2020; Odendaal 2020; Pinotti 2020; Quilty 2020; Russell TW 2020; Russell WA 2020; Ryu 2020; Shi 2020; Sruthi 2020; Steyn 2020; Taylor 2020; Utsunomiya 2020; Kwok 2020; Wells 2020; Wilson 2020; Wong MC 2020; Yang 2020; Zhang C 2020; Zhang L 2020; Zhong 2020). Modelling studies varied in the employed modelling approaches; details are presented in the 'Characteristics of included studies'.

We identified 13 observational studies assessing symptom/exposure-based screening at borders (Al-Qahtani 2020; Al-Tawfiq 2020; Arima 2020; Chen J 2020; Hoehl 2020; Kim 2020; Lagier 2020; Lio 2020; Lytras 2020; Ng 2020; Shaikh Abdul Karim 2020; Wong J 2020; Yamahata 2020). Four of these observational studies also assessed quarantine and screening at borders (Al-Qahtani 2020; Arima 2020; Chen J 2020; Shaikh Abdul Karim 2020).

Risk of bias and quality of included studies

The risk of bias (observational studies) and quality (modelling studies) of included studies is summarised in Table 1 and Table 2; these summaries are stratified by intervention type, consistent with the narrative synthesis.

We assessed risk of bias of observational studies concerned with screening or quarantining travellers using QUADAS-2. Table 1 shows that the selection of the traveller population (D1) was associated with a mix of low, unclear and high risk of bias, the index test (D2) was associated with either low or unclear risk of bias, the reference test (D3) was associated with a mix of low, unclear and high risk of bias, and the flow (D4) was generally associated with an unclear risk of bias. The population generally comprised all passengers (for studies of evacuation flights) or all travellers arriving at the airport of interest (for studies of real-world screening approaches); concerns related to instances where aspects such as symptom status determined entry into the study (e.g. febrile travellers are refused boarding), or how travellers were treated prior to the study (e.g. where travellers are quarantined prior

to travel, and symptomatic travellers are filtered out prior to the study beginning). Uncertainty around the index test existed where studies did not provide a clear threshold for categorising travellers as symptomatic. The reference test approach varied among studies, and included, for example, PCR testing upon arrival and/or symptom observation during a 14-day quarantine period with PCR testing of those who developed symptoms, PCR testing of all individuals regardless of symptom status, and/or PCR testing at the end of the quarantine period. Concerns with risk of bias were related to whether the combination of measures considered the reference test was likely to detect all infected individuals (so that individuals in asymptomatic or presymptomatic states would be discovered) or whether the intensity of the reference test was dependent on symptom status (e.g. where individuals with symptoms were tested more often and at a later stage of the quarantine period than asymptomatic individuals). Regarding the flow, an underestimation of the effect could occur if individuals were infected after the screening took place, for example during quarantine; here the risk depends on the specific quarantine facility and procedures, and these were often not well described. Individual judgements for each study can be found in Appendix 8.

Furthermore, the QUADAS-2 tool facilitates an assessment of the applicability of the studies; overall we had substantial concerns regarding the applicability of most studies, notably those conducted for evacuation flights or during a cruise ship outbreak. Thus, it is unclear how applicable the findings regarding these specific populations and screening programmes would be to more generic entry and exit screening measures aiming to screen larger numbers of travellers over an extended period of time. In contrast, we do not have major concerns regarding the applicability of findings derived from the three studies examining larger-scale screening programmes implemented indiscriminately to all arriving international travellers at an airport (Al-Qahtani 2020; Al-Tawfiq 2020; Wong J 2020). Individual judgements related to applicability for each study can be found in Appendix 8.

We appraised the quality of modelling studies using the above-described bespoke tool. Ratings for each study are found in Table 2. Studies varied widely with regards to quality, although some patterns emerged. For example, in several studies, there were concerns regarding the appropriateness of structural assumptions and input parameters (Q2 and Q4), as well as regarding an inadequate assessment of uncertainty (Q9). A major concern with the structural assumptions of a model could involve, for example, treating travel restrictions implemented in multiple countries at different times as independent of time, place and context, or making unrealistic assumptions about who travels and when they could become infected. Input parameters could be a major concern when a study assumed the sensitivity of a symptom/exposure-based entry screening measure to be 80% or of a PCR testing upon arrival to be 100%, when most empirical results suggest that these values are much lower. Major concerns with the assessment of uncertainty occurred when, for example, studies provided no assessment of whether altering the assumptions of the model influenced the results. Additionally, many studies did not conduct any validation of their models (Q5-Q8), although we did not consider this a critical flaw that would lead to 'major concerns'. Importantly, many studies did not undertake any external validation (i.e. a validation on any collected data), which we considered important with respect to the directness of the

findings (see Assessment of the certainty of evidence). Individual judgements for each study can be found in Appendix 9.

Effects of interventions/results of the synthesis

We present the effects on specific outcomes in each of the three broad outcome categories, i.e. cases avoided due to the measure, shift in epidemic development, and cases detected due to the measure.

In the following, we provide a detailed narrative summary of the impact of four broad categories of travel-related control measures.

1. Travel restrictions reducing or stopping cross-border travel (includes evidence from modelling studies only)
2. Screening at borders (includes evidence from modelling and observational studies)
3. Quarantine (includes evidence from modelling studies only)
4. Quarantine and screening at borders (includes evidence from modelling and observational studies)

For each intervention-outcome we have structured the results as follows: a full summary of findings, including a narrative summary of the effects, potential effect moderators, as well as the certainty of evidence, which can be found in the corresponding 'Summary of findings tables'. This information is also more concisely summarised in the text below. All data from the individual studies underlying these summaries can be found in the corresponding Appendices (Appendix 10; Appendix 11; Appendix 12; Appendix 13; Appendix 14).

Given that potential effect moderators were generally only assessed in individual studies (for modelling studies) or were based on limited data (for observational studies), we aimed to be cautious in our description of these below, and these data should be interpreted with caution. Although we could not explicitly assess how methodological and contextual differences across studies impacted the results, we consider these very important, and they should be kept in mind when interpreting the results described below.

1. Travel restrictions reducing or stopping cross-border travel

We identified 31 modelling studies contributing evidence to travel restrictions reducing or stopping cross-border travel, modelling different levels of reduction in travel volume. Twenty-two studies reported on the cases avoided due to the measure, 12 studies on the shift in epidemic development, and no studies on the cases detected due to the measure. A study-by-study overview of the evidence contributing to each of these outcomes is presented in Appendix 10. Summary of findings 1 presents the GRADE 'Summary of findings' for this body of evidence. While we observed a largely consistent, usually positive direction of effect, we assessed the certainty of evidence for all of these outcomes as low or very low because of risk of bias (quality), indirectness, and imprecision in the bodies of evidence.

1.1. Outcome category: cases avoided due to the measure

Number or proportion of cases in the community

Thirteen modelling studies reported on the number or proportion of cases (Anderson 2020; Banholzer 2020; Binny 2020; Chen T 2020; Costantino 2020; Deeb 2020; Kang 2020; Linka 2020a; Nowrasteh 2020; Kwok 2020; Yang 2020; Zhang C 2020; Zhong 2020) (very low-

certainty evidence). Ten of these studies reported reductions in the number or proportion of cases resulting from various travel restrictions (Anderson 2020; Banholzer 2020; Binny 2020; Chen T 2020; Costantino 2020; Deeb 2020; Kang 2020; Linka 2020a; Kwok 2020; Zhong 2020). These positive effects ranged from a 1.8% (95% confidence interval (CI) -21.9% to 17.5%) reduction in Binny 2020 to a 97.8% reduction in Kang 2020. The remaining three studies reported mixed effects, observing that a positive effect, but also no effect or even a negative effect were possible (Nowrasteh 2020; Yang 2020; Zhong 2020). Insights from specific studies highlight aspects that may influence the magnitude of effect of implementing or relaxing travel restrictions. Effects were dependent, for example, on the level of community transmission (Anderson 2020; Kwok 2020), the implementation of community-based interventions, such as a stay-at-home order, extensive testing and contact tracing (Binny 2020), and the countries restricted by the measure, with the most effective measures being those that prevented passengers from exiting regions or countries with high community transmission, such as Wuhan, China and Italy in the early stages of the pandemic (Zhong 2020).

Number or proportion of imported or exported cases

Nine modelling studies reported on the number or proportion of imported or exported cases (Adekunle 2020; Anzai 2020; Chen T 2020; Chinazzi 2020; Costantino 2020; Liebig 2020; McLure 2020; Russell TW 2020; Wells 2020) (very low-certainty evidence). Eight of these reported reductions in importations or exportations (Anzai 2020; Chen T 2020; Chinazzi 2020; Costantino 2020; Liebig 2020; McLure 2020; Russell TW 2020; Wells 2020). These positive effects ranged from an 18% reduction (Liebig 2020) to a 99% reduction (Chen T 2020) in importations or exportations. One study reported mixed effects, observing both positive effects and no effect (Adekunle 2020). Insights from specific studies suggest reasons for the observed variation in the magnitude and direction of effect. For example, earlier implementation of restrictions was shown to lead to more pronounced reductions (Liebig 2020). Travel volumes also played a role, with the proportion of countries in which imports would have contributed to over 10% of cases ranging from 56% to 75%, depending on whether flight volumes during the pandemic, in the hypothetical absence of travel restrictions, were assumed to be similar to previous years, or substantially lower (Russell TW 2020). The magnitude and direction of effect varied with the countries under study (Adekunle 2020), and the comprehensiveness and severity of the measure implemented (Costantino 2020; McLure 2020).

Number or proportion of deaths

Three modelling studies reported on the number or proportion of deaths (Binny 2020; Costantino 2020; Kwok 2020). All these studies reported reductions in deaths (very low-certainty evidence). These positive effects ranged from a 4.3% (95% CI -39.1% to 39.1%) reduction (Binny 2020) to a 98% reduction in deaths (Costantino 2020). Several aspects described in specific studies may contribute to this variation. For example, the effects were reported to depend on the presence or absence of community-based interventions, such as a stay-at-home order, extensive testing, and contact tracing (e.g. 1187 deaths when implementing quarantine of incoming travellers and border closure (except to returning residents and citizens) only and 23 deaths when implementing these interventions followed by other community-based measures (Binny 2020)). Travel restrictions implemented at

higher and lower levels of community transmission led to only a slightly different proportion of deaths avoided (14% and 12% reductions, respectively (Kwok 2020)).

Risk of importation or exportation

Three modelling studies assessed the risk of importation or exportation of cases (Nakamura 2020; Shi 2020; Zhang L 2020) (very low-certainty evidence). Two of these studies reported reductions in the risk of importing and/or exporting cases, however without providing effect estimates (Nakamura 2020; Zhang L 2020). One study reported an increased risk of importation at some airports, but decreased risk at other airports around the world as a result of loosening travel restrictions (Shi 2020). One study suggested that the country's connectedness to the international travel network and the level of community transmission are likely to play a role in the effects (Nakamura 2020).

1.2 Outcome category: shift in epidemic development

Probability of eliminating the epidemic

One modelling study assessed the probability of eliminating the epidemic (Binny 2020). The study reported mixed effects on the probability of eliminating the epidemic: the probability would be higher (66%) for border restrictions followed by strict community measures than for a delayed border closure (55% probability), and the same as early implementation of border restrictions, such as quarantine of incoming travellers (66% probability) (very low-certainty evidence). The effect of these travel restrictions were suggested to depend on the existence of community-based interventions, such as a stay-at-home order, extensive testing, and contact tracing (0% probability of eliminating the epidemic when implementing travel restrictions without community measures).

Effective reproduction number

Two modelling studies reported on changes in the effective reproduction number (R_t) (Linka 2020a; Sruthi 2020) (very low-certainty evidence). One study reported a beneficial change (i.e. break point) in R_t after the implementation of travel restrictions in European Union countries (mean duration time to the inflection point: 12.6 days) (Linka 2020a). The other study reported mixed effects (Sruthi 2020), reporting that complete border closures would lead to a 0.045 reduction in R_t, partial relaxation through the opening of land borders would lead to a 0.177 increase in R_t, while further relaxation allowing for international travel followed by quarantine upon arrival would not lead to a change in R_t.

Time to outbreak

Six modelling studies assessed the time to outbreak (Anzai 2020; Davis 2020; Grannell 2020; Linka 2020b; Odendaal 2020; Zhong 2020) (very low-certainty evidence). Four of these studies reported reductions in the time to outbreak (Anzai 2020; Davis 2020; Linka 2020b; Odendaal 2020) (very low-certainty evidence). These positive effects ranged from a delay of less than one day (Anzai 2020) to 85 days (Linka 2020b). Two studies reported mixed effects, suggesting both positive effects and no effect (Grannell 2020; Zhong 2020). In specific studies, magnitude and direction of effects were reported to depend on the presence or absence of community-based interventions and the level of community transmission (e.g. delays of 58 and 85 days for R_t=1.35 and R_t=1.16, respectively (Linka 2020b)), the timing of the implementation (e.g. travel restrictions

imposed on China implemented one week earlier would have led to an additional delay in community transmission (Davis 2020)), and the countries restricted by the measure, with the most effective measures being those that prevented passengers from exiting regions or countries with high levels of community transmission, such as Wuhan (China) and Italy in the early stages of the pandemic (Zhong 2020).

Risk of outbreak

Two modelling studies assessed the risk of an outbreak (Anzai 2020; Boldog 2020) (very low-certainty evidence). One study reported reductions in the risk of an outbreak resulting from travel restrictions with effects ranging from 1% to 37% reductions (Anzai 2020). The other study reported mixed effects, including both a positive effect and no effect (Boldog 2020). As the studies demonstrate, the variation in the magnitude and direction of effect might be explained by methodological differences between studies, as well as differences in the levels of community transmission, the number of cases in the country of departure, the severity of the travel restriction, co-interventions, and the percentage of contacts being traced. For example, larger effects were found for lower R_0 and higher proportion of contacts traced (Anzai 2020). Similarly, at lower numbers of cases in China, 25%, 50%, and 75% travel reductions resulting from restrictions implemented in Canada yielded a risk of a major outbreak of 35%, 30% and 15%, respectively; at higher numbers of cases in China, these risks were 80%, 70%, and 45%, respectively (Boldog 2020).

Number or proportion of cases at peak

Two modelling studies reported on the number of daily cases at the epidemic peak (Binny 2020; Grannell 2020). Both studies reported reductions in the number or proportion of cases at peak (low-certainty evidence). These positive effects ranged from a 0.3% reduction (Grannell 2020) to a 8% reduction (Grannell 2020). As reported in the studies, the magnitude of effect is likely to vary with the implementation of effective community-based interventions, such as a stay-at-home order, extensive testing, and contact tracing (e.g. 47,592 daily cases at peak when implementing quarantine of incoming travellers and border closure only and 80 cases when implementing these interventions followed by other community-based measures (Binny 2020).

Epidemic growth acceleration

One modelling study assessed the epidemic growth acceleration (Utsunomiya 2020). It reported that international travel controls would lead to a decrease in the growth acceleration of the epidemic progression across 62 countries (−6.05% change, P < 0.0001) (low-certainty evidence).

Exportation growth rate

One modelling study assessed the exportation growth rate (Pinotti 2020). The results suggested that both the lockdown of Hubei, resulting in a ban of all travel, as well as travel restrictions on China as a whole, led to a decrease in the growth rate of cases exported from Hubei and the rest of China to the rest of the world (low-certainty evidence).

1.3 Outcome category: cases detected due to the measure

No studies were found to contribute evidence to this outcome category.

1.4 Secondary outcomes

We identified one modelling study contributing evidence to travel restrictions reducing or stopping cross-border travel on secondary outcomes related to healthcare utilisation (Kwok 2020). This study shows that even with border closure between Hong Kong in China in place, with higher levels of community transmission (R_t = 2.2) hospitals were predicted to reach capacity by the end of March 2020. Only with low community transmission (R_t = 1.6), were hospitals predicted not to reach capacity.

2. Screening at borders

We identified 13 modelling studies contributing evidence to screening at borders, with screening modelled to reflect symptom/exposure-based screening or test-based screening and only those screening positive receiving a follow-up measure, such as self-isolation or refusal of entry. Two studies reported on the cases avoided due to the measure, four studies on the shift in epidemic development, and seven studies on the cases detected due to the measure. A study-by-study overview of the evidence contributing to each of these outcomes is presented in Appendix 11.

Additionally, 13 observational studies reported data on screening at borders. All these reported data only on cases detected due to the measure. A study-by-study overview of the evidence contributing to these outcomes, including a description of the approaches to identify cases and study data is presented in Appendix 12.

Summary of findings 2 presents the GRADE 'Summary of findings' for this body of evidence. Here we have separated bodies of evidence that reported on symptom/exposure-based screening at borders (screening for symptoms such as fever or cough and/or screening for risk factors, or when "screening" was used without further specification of the procedures) and test-based screening at borders (specifically PCR testing, when specified). While we observed a mostly consistent and positive direction of effect, we assessed the certainty of evidence for all of the outcomes as moderate (one outcome only), low, or very low because of risk of bias (quality), indirectness, and imprecision in the bodies of evidence.

2.1. Outcome category: cases avoided due to the measure

Symptom/exposure-based screening at borders

We identified one modelling study assessing the impact of symptom/exposure-based screening at borders on the cases avoided due to the measure (Wells 2020).

Proportion of cases exported

One modelling study assessed the number or proportion of cases exported (Wells 2020). The results suggested that putting screening measures in place across the world would reduce the number of cases exported per day from China would be reduced by 82% (95% CI 72% to 95%), under the assumption of only 35.7% of symptomatic individuals being detected (moderate-certainty evidence).

Test-based screening at borders

We identified one modelling study assessing the impact of test-based screening at borders on the cases avoided due to the measure (Dickens 2020).

Proportion of secondary cases

One modelling study examined the proportion of secondary cases due to international travel (Dickens 2020). PCR testing all incoming travellers upon arrival, followed by isolation of test-positives and requiring a negative test at the end of the isolation would lead to a reduction in secondary cases of 88% (95% CI 87% to 89%) for a 7-day isolation period and 92% (95% CI 92% to 93%) for a 14-day isolation period (very low-certainty evidence).

Proportion of imported cases

One modelling study assessed the proportion of imported cases (Dickens 2020). PCR testing all incoming travellers upon arrival, followed by isolation of test-positives and requiring a negative test at the end of the isolation would lead to a reduction of 90% of imported cases for a 7-day isolation period and 92% for a 14-day isolation period (very low-certainty evidence). Testing all incoming travellers and refusing entry to test-positives would lead to a reduction of 77%.

2.2 Outcome category: shift in epidemic development

Symptom/exposure-based screening at borders

We identified four modelling studies assessing the impact of symptom/exposure-based screening at borders on the shift in epidemic development (Clifford 2020a; Mandal 2020; Nuckchady 2020; Wilson 2020).

Time to outbreak

Four modelling studies assessed time to outbreak (Clifford 2020a; Mandal 2020; Nuckchady 2020; Wilson 2020). All studies reported that entry and/or exit screening alone would delay an outbreak (very low-certainty evidence). These positive effects ranged from 2.7-day delay (from 45 days to 47.7 days in reaching 1000 cases) (Mandal 2020) to 0.5-year delay (from 1.7 years (95% CI 0.04 to 6.09) to 2.2 years (95% CI 0.6 to 8.11)) (Wilson 2020). Insights from specific studies highlight aspects that may influence the magnitude of effect of entry and/or exit screening. For example, effects were reported to depend on the timing of the implementation (Clifford 2020a), the number of incoming travellers (Wilson 2020), the percentage of asymptomatic travellers screened (Mandal 2020), and the sensitivity of the screening (e.g. entry or exit screening with a sensitivity of 64% would delay an outbreak by 9.7 days, while screening with a sensitivity of 100% would delay an outbreak by 20 days (Nuckchady 2020)).

Risk of outbreak

One modelling study assessed the risk of outbreak (Nuckchady 2020). The results suggested that under the assumption of one infected person entering Mauritius per 100 days, entry screening with 100% sensitivity would reduce the probability of an outbreak within 3 months to 10% and screening with 50% sensitivity would reduce the probability to 48% (low-certainty evidence).

Test-based screening at borders

We did not identify any study assessing the impact of test-based screening at borders on the shift in epidemic development.

2.3 Outcome category: cases detected due to the measure

Symptom/exposure-based screening at borders

We identified four modelling studies (Bays 2020; Gostic 2020; Quilty 2020; Taylor 2020) and nine observational studies (Al-Qahtani 2020; Arima 2020; Chen J 2020; Hoehl 2020; Kim 2020; Lytras 2020; Ng 2020; Wong J 2020; Yamahata 2020) assessing the impact of symptom/exposure-based screening at borders on the cases detected due to the measure.

Number or proportion of cases detected

Four modelling studies reported on the number or proportion of cases detected (Bays 2020; Gostic 2020; Quilty 2020; Taylor 2020). All studies reported reductions in the number or proportion of cases detected (very low-certainty evidence). These positive effects ranged from detecting 0.8% (95% CI 0.2% to 1.6%) of cases (Taylor 2020) to detecting 53% (95% CI 35% to 72%) cases (Quilty 2020). Insights from specific studies suggest relevant sources of variation in the magnitude of effect. For example, the number or proportion of cases detected was reported to be influenced by the time window in which the exposure may have occurred and the duration of the flight with longer flights increasing the likelihood that symptoms develop during the flight and thus are detected (Bays 2020). The effects were also reported to depend on the percentage of asymptomatic cases in the population (Gostic 2020), the combination of entry and exit screening measures (Gostic 2020; Quilty 2020), and the sensitivity of screening (e.g. assuming a sensitivity of 86% for thermal scanner-based screening and 17% of asymptomatic cases being undetectable, entry and exit screening combined and entry screening alone would both detect 53% (95% CI 35% to 72%) of cases (Quilty 2020)).

Proportion of cases detected and positive predictive value

Nine observational studies provided data on symptom/exposure-based screening at borders (e.g. focused on the presence of fever and/or cough and/or shortness of breath) (Al-Qahtani 2020; Arima 2020; Chen J 2020; Hoehl 2020; Kim 2020; Lytras 2020; Ng 2020; Wong J 2020; Yamahata 2020); each of these measures also involved subsequently quarantining all travellers for fourteen days, independent of whether these were screened positively or negatively, usually with some form of symptom observation and sometimes further testing. For this body of evidence on symptom/exposure-based screening at borders, however, this quarantine period is not considered part of the intervention, but instead serves as a way to identify the 'true' number of cases in the study population; although even with these features, false positives and false negatives remain possible. These studies reported data, which allowed for the calculation of the proportion of cases detected by the measure. Six of these studies also reported data allowing for the calculation of the positive predictive value (PPV) (Arima 2020; Hoehl 2020; Kim 2020; Lytras 2020; Ng 2020; Yamahata 2020).

The proportion of cases detected by the screening measure varied widely (very low-certainty evidence). One study reported that the measure detected 100% of cases (Kim 2020); this was, however, an outlier, with the rest of studies reporting substantially lower proportions of cases detected. The proportion of cases detected by symptom/exposure screening is summarised in Figure 2 (top panel). The PPV, calculated only for studies assessing symptom/exposure screening, also varied widely between studies (very low-certainty evidence).

Figure 2. **Summary of the proportions of cases detected by the measure from observational studies. Measures portrayed include exit and/or entry screening (top panel) and PCR tests (middle panel), as well as for combined measures exit and/or entry screening with quarantine and further screening, in the form of symptom observation and/or PCR tests (bottom panel). Notes: Yamahata 2020 employed a form of symptom screening aboard a cruise ship, thus representing a very different context than all other studies. Ng 2020 employed a delayed PCR test on day 3. Lagier 2020 and Lio 2020 employed a PCR test on arrival and on day 2, respectively, however given that they did not identify cases they are not portrayed in this figure. The five evacuation flights assessed in Shaikh Abdul Karim 2020 had very different COVID-19 prevalences, with no cases associated with three flights, but with 2/104 and 80/124 on the remaining two flights.**
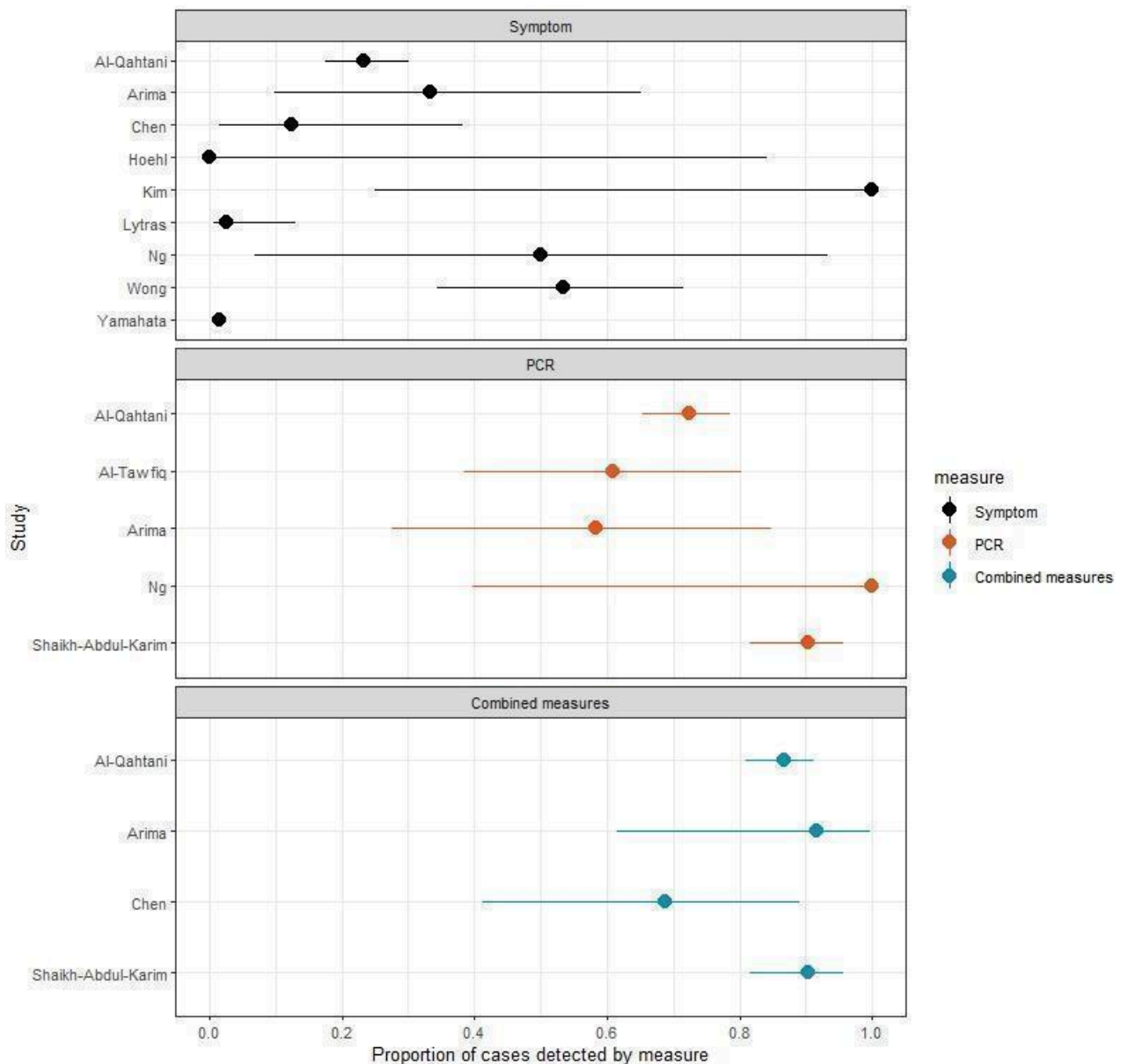
The individual screening measures themselves and the context in which they were implemented are important aspects to consider in interpreting these results. Among the symptom/exposure screening measures, the screening approaches (e.g. screening for fever, for any kind of respiratory symptoms and/or for contact with COVID-19 cases in the past days with these measures being performed prior to departure, upon arrival, or both), as well as the approaches for determining cases vary across studies, and for many studies it was unclear what threshold was used for determining whether an individual was symptomatic. Most studies reported on measures implemented in very specific settings, i.e. either as part of evacuation flights or on a cruise ship, while only two studies assessed national-level border control measures (Al-Qahtani 2020; Al-Tawfiq 2020).

Test-based screening at borders

We identified three modelling (Clifford 2020b; Russell WA 2020; Steyn 2020) studies and five observational studies (Al-Qahtani 2020; Al-Tawfiq 2020; Arima 2020; Ng 2020; Shaikh Abdul Karim 2020) assessing the impact of entry and/or exit test-based screening on the cases detected due to the measure.

Days at risk of transmitting the infection into community

Two modelling studies reported on the days that travellers, upon release, remain at risk of transmitting the infection into the community (Clifford 2020b; Russell WA 2020). Both studies reported that a single PCR test upon arrival would reduce the days at risk of transmission (low-certainty evidence). These positive effects ranged from 0.1 fewer days (Clifford 2020b) to 0.3 fewer days at risk of transmission (Russell WA 2020).

Probability of releasing an infected individual into the community

Two modelling studies reported on the probability of releasing an infected individual into the community (Clifford 2020b; Steyn 2020). Both studies reported reductions in the probability of releasing an infected individual into the community as a result of PCR testing (low-certainty evidence). These positive effects included a risk ratio of 0.55 (95% CI 0.28 to 0.83) (Clifford 2020b) and probabilities of releasing an infected individual ranging from 48% to 53% for scenarios with different risks of transmission while travelling (Steyn 2020).

Proportion of cases detected and positive predictive value

Five observational studies provided data on PCR testing (Al-Qahtani 2020; Al-Tawfiq 2020; Arima 2020; Ng 2020; Shaikh Abdul Karim 2020); four studies conducted the test within 24 hours, one study after a delay of three days (Ng 2020). Each measure also involved subsequently quarantining all travellers for fourteen days, independent of whether these tested positively or negatively, usually with some form of symptom observation and a test for all individuals at the end of the quarantine period. As described for the symptom/exposure screening above, for this body of evidence this quarantine period is not considered part of the intervention, but instead serves as a way to identify the 'true' number of cases in the study population, although even with these features, false positives and false negatives remain possible. These studies reported data, which allowed for the calculation of the proportion of cases detected by the measure; two further studies (Lagier 2020; Lio 2020), which conducted tests within 24 hours and two days after arrival, respectively, identified no cases, meaning that we could not report the proportion of cases detected.

The proportion of cases detected by testing varied (58.3% to 90.2%) (low-certainty evidence). The proportion of cases detected by testing is summarised in Figure 2 (middle panel). The PPV was not calculated for studies assessing PCR testing, as those with a positive PCR test at a given point were considered true cases; no data were available to determine false positives.

The individual testing measures themselves and the context in which they were implemented are important aspects to consider in interpreting these results. For example, the proportions of cases detected for Al-Qahtani 2020 and Shaikh Abdul Karim 2020 were 58.3% and 90.2%, respectively. The prevalences differed, however, with 188 of 2714 (6.9%) and 82 of 432 (19.0%), respectively, being infected. Looking further, Shaikh Abdul Karim 2020 examined five

flights; no cases were identified for three of the flights, while 2 of 104 (2.0%) and 80 of 124 (65.0%) on the remaining two flights. The screening approaches varied somewhat, for example with respect to timing of test provision; most of the studies tested within the first 24 hours, while one study tested on day 3 (Ng 2020). Most studies reported on measures implemented in very specific settings, i.e. as part of evacuation flights, while only two studies assessed national-level border control measures (Al-Qahtani 2020; Al-Tawfiq 2020).

3. Quarantine

We identified 12 modelling studies assessing the quarantine of travellers alone, comprising voluntary or government-mandated quarantine of travellers of different duration without any accompanying or follow-up measures. Six studies reported on cases avoided due to the measure, one study on the shift in epidemic development, and five studies on the cases detected due to the measure. A study-by-study overview of the evidence contributing to each of these outcomes is presented in Appendix 13. Summary of findings 3 presents the GRADE summary of findings for this body of evidence. While we observed a consistent, largely positive direction of effect, we assessed the certainty of evidence for all of the outcomes as low or very low because of risk of bias (quality), indirectness, and imprecision in the bodies of evidence.

3.1 Outcome category: cases avoided due to the measure

Number or proportion of cases in the community

Three modelling studies examined the number or proportion of cases (Chen T 2020; Chen Y-H 2020; Wong MC 2020). All studies reported reductions in the number or proportion of cases (very low-certainty evidence). These positive effects ranged from 450 fewer cases in Wong MC 2020 to 64,028 fewer cases in Chen T 2020 during the first wave of the pandemic. Insights from specific studies suggest that the effects might depend on the target group (e.g. quarantining all inbound travellers versus only those that are symptomatic, with the former predicting larger reductions in the number of cases (Chen T 2020)).

Proportion of imported cases

One modelling study assessed the proportion of imported cases (Dickens 2020). The study reported that quarantining all incoming travellers would reduce the proportion of imported cases by 55% for a 7-day quarantine period and by 91% for a 14-day quarantine period (very low-certainty evidence).

Number or proportion of cases seeded by imported cases

Three modelling studies reported on the number of cases seeded by imported cases (Dickens 2020; James 2020; Ryu 2020). All studies reported reductions in the number or proportion of cases seeded by imported cases as a result of quarantine of travellers (very low-certainty evidence). These positive effects in James 2020 ranged from 26% (95% CI 19% to 37%) reduction to 100% (95% CI 62% to 100%) reduction. Reductions were larger when the quarantine was government-mandated (James 2020), for the elderly compared with adults (James 2020), and for longer quarantine periods (Dickens 2020).

Probability of an imported case not infecting anyone

One modelling study assessed the probability of an imported case not causing further infections (James 2020). The study reported that a 14-day self-isolation of all international arrivals in New

Zealand would lead to 4% and 14% increase in the probability that an imported case would not infect anyone among adults and the elderly, respectively. The increase in the probability would be higher when a 14-day government-mandated quarantine is required (31% and 36% among adults and the elderly, respectively).

3.2 Outcome category: shift in epidemic development

Time to outbreak

One modelling study reported on the time to outbreak (Kivuti-Bito 2020). The study reported that increasing the effectiveness of quarantine of travellers to 80% and 90% from the base case of 75% effectiveness would delay the peak in active cases and deaths by 3.5 and 5.5 days, respectively (low-certainty evidence).

3.3 Outcome category: cases detected due to the measure

Days at risk of transmitting the infection into community

Two modelling studies assessed the days that travellers will be at risk of transmitting the infection into the community (Clifford 2020b; Russell WA 2020). Both studies reported reductions in the numbers of days at risk of transmission resulting from quarantine (low-certainty evidence). These positive effects ranged from 0.1 fewer days (Clifford 2020b) to 2.1 fewer days at risk (Clifford 2020b). The studies reported that the variation in the magnitude of effect might be explained by the length of quarantine with longer quarantine periods predicting larger effect (e.g. 2-day quarantine: 1.8 days at risk (95% CI 1.6 to 2.2); 14-day quarantine: 0.53 days at risk (95% CI 0.46 to 0.60) (Russell WA 2020)).

Proportion of cases detected

One modelling study examined the proportion of cases detected (Taylor 2020). The study reported that requiring travellers to quarantine upon arrival in the UK would lead to detecting different proportion of cases, with the magnitude increasing with the number of days in quarantine (7-day quarantine: 51% (95% CI 47% to 56%); 14-day quarantine: 78% (95% CI 74% to 82%)) (very low-certainty evidence). These proportions are higher than those for screening alone (with either thermal imaging scanners or health checks detecting 0.78% and 1.13% of cases, respectively).

Probability of releasing an infected individual into the community

Three modelling studies examined the probability of releasing an infected individual into the community (Ashcroft 2020; Clifford 2020b; Steyn 2020). All studies reported reductions in the risk or probability of releasing an infected individual (very low-certainty evidence). These positive effects included a risk ratio ranging from 0.00 (95% CI 0.00 to 0.01) to 0.59 (95% CI 0.28 to 0.85) (Clifford 2020b) and probabilities of releasing an infected individual ranging from 0% (Steyn 2020) to 85% (Ashcroft 2020). Insights from these studies suggest that the magnitude of effects might depend on the length of the quarantine period (Clifford 2020b), duration of travel (Ashcroft 2020), and the risk of transmission within quarantine settings (Steyn 2020).

3.4 Secondary outcomes

We identified three modelling studies contributing evidence to quarantine only on secondary outcomes related to infectious disease transmission (Ashcroft 2020; Chen Y-H 2020; Steyn 2020).

One study assessed quarantine utility as a ratio of the amount of transmission prevented to the number of person days spent in quarantine, predicting that for long-duration travel (i.e. 7 days or longer), shorter quarantine provides a better balance between preventing infection and the burden of quarantining individuals than a longer quarantine duration. For short duration travel, however, this relationship is reversed (Ashcroft 2020). One study reported that with test-and-isolation, contact tracing, and general public mask-wearing and other social measures in place, strict quarantine of travellers (one daily infection imported) in Taiwan would ensure that the number of individuals needed to quarantine in the community remains low (4092) over 90 days. Without quarantine (10 daily infections imported), the number of individuals needed to quarantine in the community would increase steadily (40810) over the same time period (Chen Y-H 2020). One study suggests that there is a risk that front-line workers from quarantine facilities seed outbreaks in the community. With weekly testing of front-line workers from quarantine facilities, there is a high probability that a case will be detected in the front-line worker as opposed to later in a secondary case in the community. With less frequent or no testing of front-line workers, the probability increases that the case is not detected until a secondary case is infected (Steyn 2020).

4. Quarantine and screening at borders

We identified seven modelling studies assessing quarantine of travellers and screening at borders and/or at different days during quarantine. No studies reported on the cases avoided due to the measure, one study reported on the shift in epidemic development, and six studies reported on the cases detected due to the measure. A study-by-study overview of the evidence contributing to each of these outcomes is presented in Appendix 14.

Additionally, we found four observational studies contributing data to this intervention category, all reporting on the cases detected due to the measure. A study-by-study overview of this evidence, including a description of the approaches to identify cases and study data is presented in Appendix 12.

Summary of findings 4 presents the GRADE summary of findings for this body of evidence. While we observed a consistent and positive direction of effect, we assessed the certainty of evidence for all of these outcomes as low or very low because of risk of bias (quality), indirectness, and imprecision in the bodies of evidence.

4.1 Outcome category: cases avoided due to the measure

No studies were found to contribute evidence to this outcome category.

4.2 Outcome category: shift in epidemic development

Time to outbreak

One modelling study reported on the time to outbreak (Wilson 2020). The study reported delays in outbreak resulting from quarantine and screening at borders measures. Under the assumption of one flight per day (7.1% of normal travel volume) and 50% sensitivity, the time to outbreak would vary greatly for different combinations of measures ranging from 3.5 years (95% CI 0.09 to 12.9) to 34.1 years (95% CI 0.86 to 126) to outbreak (very low-certainty evidence). Combination of measures, such as exit screening, in-flight wearing of masks, PCR testing of arriving travellers and quarantine would lead to larger delays of outbreak.

4.3 Outcome category: cases detected due to the measure

Days at risk of transmitting the infection into community

Two modelling studies assessed days that the travellers remain at risk of transmitting the infection into the community (Clifford 2020b; Russell WA 2020). Both studies reported that the combination of quarantine and testing would reduce days at risk of transmission compared with single measures (low-certainty evidence). These positive effects ranged from 0.01 fewer days (Russell WA 2020) to 2.0 fewer days (Clifford 2020b) at risk of transmission. Requiring two tests before releasing from quarantine showed slightly improved, yet largely comparable effects to the effects of quarantine with a single test at the end (Clifford 2020b).

Probability of releasing an infected individual into the community

Three modelling studies assessed the probability of releasing an infected individual into the community (Ashcroft 2020; Clifford 2020b; Steyn 2020) (very low-certainty evidence). All three studies reported positive effects resulting from a combination of screening and quarantine. These positive effects included a reduction in the probability of releasing an infected individual ranging from 2% to 48% (Steyn 2020). Studies reported that the variation in the magnitude of effect could be explained by the length of quarantine, with shorter periods predicting smaller effects (e.g. risk ratio (RR) for a 3-day quarantine with a single test: 0.22 (95% CI 0.02 to 0.48) and RR for a 14-day quarantine with a single test: 0.01 (95% CI 0.00 to 0.03)) (Clifford 2020b), the single measure comparison used (e.g. quarantine alone or PCR testing upon entry alone), day(s) on which the test is administered (Ashcroft 2020), or the risk of transmission within quarantine (Steyn 2020).

Proportion of cases detected

Two modelling studies reported on the proportion of cases detected (Bays 2020; Taylor 2020). Both studies reported that the combination of quarantine and testing would further increase case detection compared with single measures (very low-certainty evidence). The positive effects ranged from 41% (Bays 2020) to 99% of cases detected (Bays 2020). Adding a second test suggested only a slight improvement in the effect (Taylor 2020). Insights from specific studies suggest that the observed variation in the magnitude of effect may be explained by the length of the quarantine period with longer quarantine increasing the detection rate (Bays 2020; Taylor 2020), and the duration of travel and stay in the country of departure (Bays 2020).

Proportion of cases detected and PPV

Four observational studies provided data on quarantine and screening at borders (Al-Qahtani 2020; Arima 2020; Chen J 2020; Shaikh Abdul Karim 2020). Each of these studies began with an entry and/or exit screening measure, comprising symptom screening in one study (Chen J 2020), and PCR testing upon arrival in the other three. Each measure also involved subsequently quarantining all travellers for fourteen days; all but one study (Shaikh Abdul Karim 2020), additionally monitored symptoms of all travellers and tested those who developed symptoms. Above, we considered this quarantine period as part of the way to identify the 'true' number of cases; for this body of evidence, however, we consider this as a 'combined measures' intervention. Each study also provided PCR testing before release from quarantine; we treat this final PCR test as the way to identify the 'true' number of cases.

These studies reported data, which allowed for the calculation of the proportion of cases detected by the measure.

The proportion of cases detected by these combined measures, comprising quarantine of travellers and screening, is summarised in Figure 2 (bottom panel). As visible in the figure, across the studies, in comparison to exit and/or entry screening only, a subsequent 14-day quarantine period with symptom monitoring and further testing led to the detection of additional cases that would have been missed by the initial screening measure. Only one study did not detect further cases after the initial entry and/or exit screening measure (Shaikh Abdul Karim 2020) (low-certainty evidence). The PPV was not calculated for studies assessing PCR testing, as those with a positive PCR test at a given point were considered true cases; no data were available to determine false positives.

The individual measures themselves and the context in which they were implemented are important aspects to consider in interpreting these results. As described above, one study employed symptom screening (Chen J 2020), while the others used PCR testing upon arrival. Measures employed during the quarantine itself differed as well; while most studies monitored for the development of symptoms and tested those developing symptoms, one study did not (Al-Qahtani 2020); the intensity with which symptoms were monitored could also be important. Three studies reported on measures implemented in very specific settings, i.e. as part of evacuation flights, while only one study assessed a national-level border control measure.

DISCUSSION

Summary of main findings

To inform decisions on containing the COVID-19 pandemic, we updated our previous review (Burns 2020), with the aim of identifying and synthesising the evidence on the effectiveness of international travel-related control measures during coronavirus outbreaks on infectious disease transmission and screening-related outcomes. We identified a much expanded and heterogeneous evidence base, with studies focusing on a range of real or simulated travel-related control measures aiming to contain the COVID-19 pandemic. In the original review we assessed studies on SARS and MERS (Burns 2020); in this update, we focus on COVID-19 studies.

We found 31 modelling studies on travel restrictions reducing cross-border travel, all modelling a range of reductions in travel across real or simulated countries. Studies reported on various outcomes related to cases avoided due to the measure and shift in epidemic development. Across outcomes, most studies predicted a positive effect; some studies, however, observed mixed effects, including positive and negative effects. Very low- to moderate-certainty evidence limits our confidence in these findings.

We found 13 modelling studies and 13 observational studies on screening at borders. Screening measures covered symptom/exposure-based screening and/or PCR test-based screening before departure or upon or soon after arrival. Regarding symptom/exposure-based screening at borders, modelling studies assessed several different outcomes related to shift in epidemic development and cases detected due to the measure; observational studies assessed outcomes related to cases detected

Cochrane
Library

Trusted evidence.
Informed decisions.
Better health.

Cochrane Database of Systematic Reviews

due to the measure. For all outcomes, the observed findings showed positive effects, although some of these effects were very small; effects were dependent on factors, such as the sensitivity of screening measures. Regarding test-based screening at borders, modelling studies assessed outcomes related to cases avoided and cases detected due to the measure; observational studies assessed cases detected due to the measure. Across these outcomes, the findings showed positive effects with magnitudes varying, depending, for example, on the timing of testing. Although a wide range of positive effects was observed, these were generally larger than for symptom/exposure-based screening alone. Very low to moderate-certainty evidence limits our confidence in the findings on screening measures.

We found 12 modelling studies on quarantine. Studies assessed multiple outcomes related to cases avoided due to the measure, shift in epidemic development and cases detected due to the measure. Included studies all showed positive effects ranging from small to large in magnitude, depending on the quarantine duration and compliance. Very low- to low-certainty evidence limits our confidence in these findings.

We found seven modelling studies and four observational studies on quarantine and screening at borders. Studies assessed outcomes related to shift in epidemic development and cases detected. Most studies showed positive effects for the combined measures with varying magnitudes of effect depending on how the measures were combined (e.g. the length of the quarantine period and days when the test was conducted in quarantine). Very low- to low-certainty evidence limits our confidence in these findings.

Overall completeness and applicability of the evidence

Consistent with the original review (Burns 2020), we identified studies assessing a broad range of travel-related control measures to contain the COVID-19 pandemic. They examined outcomes across all three a priori specified outcome categories, and were conducted across multiple world regions. There were, however, some gaps in the evidence base, notably in relation to populations, settings and interventions.

Population

Modelling studies across all categories of travel-related control measures generally considered nonspecific populations, using data both observed in and/or modelled upon a travelling population or the general population. Observational studies of screening at borders measures used data observed from travelling populations. Many studies assessed or modelled all modes of travel or did not specify the mode (29 studies); air travellers (33 studies) were much more frequently represented than those travelling by ship (one study) or by land (one study).

Setting

We identified studies from all world regions. In contrast to the original review, this update also found studies from the African and Eastern Mediterranean regions; however, these make up only a small share of the evidence base (one study and three studies, respectively). Moreover, the screening at borders measures assessed here were largely implemented in very specific settings, such as on evacuation flights or during cruise ship outbreaks. We did, however, further identify three studies from Bahrain, Brunei and Saudi Arabia evaluating measures at a population-

level, such as all travellers arriving at a specific airport or a large population of workers returning from work abroad. While likely more policy-relevant than those in very specific settings, the populations studied were small, in the order of a few thousand, and an opportunity for undertaking much larger and thus more policy-relevant studies exists. Importantly, much of the evidence relates to the implementation of travel-related control measures at the beginning of the COVID-19 pandemic, although some studies were conducted during later phases.

Intervention

With much of the evidence deriving from modelling studies – notably for travel restrictions reducing cross-border travel and quarantine of travellers – there is a lack of 'real-world' evidence for many of these measures. Compared to the original review, however, we identified modelling studies on screening at borders measures that more closely matched the current policy discussions. For example, whereas earlier studies asked very generic questions such as 'Is screening effective at detecting infected travellers?', some of the more recent studies asked more nuanced questions such as 'Does a PCR test upon arrival perform better than symptom/exposure screening?' and 'How many more cases are detected if a PCR test is given after a quarantine of 3, 7 or 14 days?'. Consistent with the original review, little evidence was found on the relaxation of travel-related control measures; as various countries consider when it is safe to lift restrictions, it will be important that studies assessing this aspect are conducted.

Outcomes

Within our primary outcome categories comprising infectious disease transmission and screening-related outcomes, studies assessed a range of outcomes; notably studies identified in the update addressed a larger number of outcomes. In the original review, studies assessing entry and/or exit screening and quarantine of travellers focused only on the proportion of infected individuals detected by the measure; newer studies also assessed the number of infectious individuals released into the community and the amount of time an individual is infectious after arrival and after being quarantined for different durations. Regarding secondary outcomes, few studies reported on the impact of travel-related control measures on the number of cases seeded by frontline workers, the number of hospitalisations and the number of individuals needing to quarantine/isolate in the community. While it is possible that travel-related control measures generate signalling effects in terms of raising general awareness of the risk of infection or deterring effects in terms of stopping sick individuals from travelling, we did not identify any evidence on these outcomes.

No studies included in this update assessed outcomes concerned with the human and financial resources required to implement the measures or adverse effects in terms of health (e.g. isolation), as well as broader social and economic implications (e.g. stigmatisation, inability to work, economic impacts). This represents a major limitation regarding the completeness of the evidence, as this information is important to assess the benefits and harms of the measures. At the outset, it was decided in consultation with the WHO that the most pressing question at this point was the effectiveness of travel-related control measures and that this review should thus focus primarily on studies assessing effectiveness in relation to infectious disease transmission and

screening-related outcomes; if included studies also reported on harms, these data would also be examined. In addition, a separate, currently ongoing, scoping review of the health, social sciences, and environmental literature seeks to map the various unintended consequences and potential adverse health effects and broader societal harms of travel-related control measures (osf.io/7gyxe).

It is important to note that this is a fast-moving research field. Since 13 November 2020, when we conducted our searches, we have identified two further relevant studies: a modelling study on border screening approaches using various testing strategies in the USA (Kiang 2020), and an observational study of test-based screening at borders using PCR testing in New Zealand (Swadi 2021). As these studies were not identified as part of our systematic searches and were published after the final searches, we have not incorporated them into this review. They highlight, however, that the evidence base is growing further, and that a future update will be important.

Sources of heterogeneity

As part of the narrative synthesis, we documented potential sources of heterogeneity that may have influenced intervention effectiveness. Modelling studies across all intervention categories differed in the methods they employed and they assessed a broad range of potential factors.

- COVID-19 pandemic: studies suggest that the level of community transmission in both the implementing country and the restricted country, and the proportion of asymptomatic cases play a role.
- Broader context of travel: the baseline number of travellers, the interconnectedness of the region with the travel measure in place and the restricted region, how much flight volumes are likely to rebound in the absence of restrictions, the duration the duration of travel and stay in the country of departure were all relevant factors.
- Other public health measures: whether other public health measures, such as a stay-at-home order and testing and contact tracing, are in place in the region where the travel measure is implemented.
- Implementation of the intervention: factors related to the earlier or later timing of implementation of the intervention, the exact specification of the intervention (e.g. duration of quarantine of travellers), and compliance with the measures all influenced effectiveness.

Looking across the observational studies of screening at borders, the measure itself was important, with test-based screening generally performing better than symptom- or exposure-based screening. There were also some differences regarding the timing of the interventions; while most measures were implemented at departure or immediately upon arrival, a few studies assessed testing within one or two days of arrival.

Certainty of the evidence

The certainty of evidence was moderate for one outcome and either low or very low for the rest of the body of evidence we assessed; thus, we cannot be confident in the findings. The true effects may therefore be (or are likely to be) substantially different from the estimates of effect described. We downgraded the certainty of evidence due to risk of bias (for observational studies) and quality (for modelling studies), as well as for imprecision and indirectness.

Observational studies contributed evidence to the intervention categories entry and/or exit screening alone and quarantine of travellers alone. We judged most domains across observational studies assessing entry and/or exit screening measures to be at low or unclear risk of bias. There were, however, some studies at high risk of bias for the selection of travellers, the reference test and the flow (where, for example, travellers were potentially infected while in quarantine, after the screening took place).

Modelling studies contributed evidence to all four intervention categories. Although modelling studies differed in quality, most bodies of evidence comprising modelling studies were downgraded due to serious concerns about the quality of the modelling. Quality concerns were diverse, but included inappropriate or unrealistic assumptions related to model structure and input data, the lack of assessment of uncertainty and incomplete technical documentation. Problematic assumptions for any of these aspects could lead to results that do not reflect reality and are thus of limited utility.

We used four reasons for downgrading evidence based on imprecision. We downgraded evidence for imprecision when a body of evidence comprised a single modelling study, as it limits our confidence in the predictions being a precise estimate of true effects, or when multiple studies provided a wide range of plausible effects (e.g. no effect versus large reductions in the number of cases). Furthermore, a few of the modelling studies provided no estimates of effect (e.g. data presented in a diagram or a map), and many studies provided estimates of effect (e.g. 85 deaths avoided) with insufficient information on the precision of the estimates. Given the nature of the data and models, it is plausible that the uncertainty in estimates is wide, and such information would be necessary for an appropriate interpretation of the study findings.

In this update, we applied two reasons for downgrading based on indirectness. Where exit and entry screening measures were implemented in very specific settings, such as on evacuation flights or during cruise ship outbreaks, we considered this as indirect evidence with regards to informing more general entry and exit screening measures at national borders. Additionally, for bodies of evidence based on modelling studies, we downgraded evidence for indirectness when there was no external validation of the model(s), as it created uncertainties in assessing how directly the model outputs relate to real-world outcomes and consequently to our review question. External validation may be challenging, especially in the context of a pandemic, but it is important to ensure that the findings are generalisable to the real-world situation.

Potential biases in the review process

In this update, consistent with the original review, we applied systematic and transparent methods throughout the phases of the review process. We defined our review objective and scope informed by a previously conducted evidence map (Movsisyan 2021), and in consultation with the WHO – a key end-user that specifically requested the review to inform WHO recommendations on travel-related control measures for COVID-19. Our protocol was reviewed and approved by Cochrane (see Appendix 1). In order to describe the emerging evidence in relation to COVID-19, we included a wide range of study designs and publication types, including modelling studies and preprint publications. We synthesised this evidence narratively, but applied GRADE to assess the certainty of evidence for all primary outcomes. We did, however,

encounter challenges in dealing with this complex evidence base, and some decisions we made may have introduced bias into the review process.

Although we used a comprehensive search strategy designed by an information specialist, we conducted searches in only two major databases and two COVID-specific databases and used specific search terms tightly defined around travel-related control measures. While our chosen sources include records from a wide range of databases, as well as grey literature, such as preprints, it is possible that our searches missed some studies, especially if these were not appropriately indexed in the journals and preprint servers or conducted in languages other than English (e.g. Chinese literature). This also concerns our lack of findings regarding implementation outcomes and adverse effects. Had we searched a broader range of multi-disciplinary databases and had we undertaken systematic searches of the grey literature (e.g. government reports), we might have identified some of these outcomes – albeit at the cost of a much lengthier and more complex review process.

As described above, our review included many modelling studies: in the specific context of a global pandemic, models developed to make predictions about the future often represent the only available evidence and are therefore crucial in informing decision making. Many modelling studies did not provide comprehensive reporting of key assumptions and model parameters, which created challenges in assessing their eligibility and validity, for example in decisions on whether the model used disease parameters of relevance for our review. Given the lack of a validated tool to assess the quality of modelling studies, we had developed a bespoke quality appraisal tool and implemented two post-protocol adaptations in the original review. We used the same tool for the present update, and also included modelling experts within our review team throughout the review process.

We applied a structured method for the narrative synthesis that relied on defining the direction of effect for each individual study, drawing on recent guidance for conducting synthesis without meta-analysis (Campbell 2020; Hilton Boon 2020). We described this method clearly, applied it consistently across all studies, and reported the results consistently for all bodies of evidence. Due to the nature of the outcomes, our consideration of any effect greater than the null effect being potentially relevant, and the analytical methods applied in the included studies, however, this method may bias the results towards a positive effect. A study that evaluates the proportion of infected travellers detected by a screening measure, for example, will always detect some proportion of infected travellers greater than zero; this means that a screening measure detecting 1% of cases would be considered a 'positive' effect, as it detects a higher proportion of cases than would be detected with no measure in place (0%). By reporting effect ranges and providing the underlying data, however, we have aimed to be clear and transparent that some 'positive' effects are very small.

Further, we experienced some difficulty with applying the GRADE guidance for assessing the certainty of evidence based on modelling studies (Brozek 2021). Most importantly, because it does not offer guidance for operationalising the assessment of risk of bias/quality, indirectness, imprecision, inconsistency and publication bias for a body of evidence comprising multiple models. Notably, applying the criteria of inconsistency

and imprecision were challenging. With inconsistency, it was challenging because travel-related control measures by design generally show at least a slight positive or no effect, not a negative effect. With imprecision, it was challenging because high-quality models vary and use a large number of parameters and scenarios, often leading to wide confidence intervals; poor-quality models do not even allow for an assessment of imprecision due to lack of reporting of effect estimates or confidence intervals. Furthermore, we used external model validation as a key criterion to help in our judgments on indirectness of the evidence, which is, however, not currently specified and operationalised as such in the GRADE guidance on modelling studies. Finally, it should be noted that there are simply more opportunities for larger bodies of evidence to be downgraded than those with only a small number of contributing studies, or even only one contributing study, as additional studies were likely to contribute further issues on risk of bias, indirectness, and imprecision to the body of evidence. Thus, a body of evidence with one study, for example, could potentially be assessed as moderate-certainty evidence, while a body of evidence with 13 contributing studies had very little chance of being assessed as higher than very low-certainty evidence.

While we made a case in the review for the methodological and contextual differences across the studies to impact the results and for their consideration when interpreting the review findings, we were not able to formally assess these potential moderators through subgroup analyses. Our statements regarding these moderators were therefore largely based on their assessment in individual modelling studies or limited data from observational studies.

Finally, we used abridged procedures of systematic reviewing at certain stages, to enable rapid completion of this review. Specifically, we did not conduct double data extraction and assessment of risk of bias or quality appraisal. However, we had a second experienced review author check all the extracted and appraised data, and discussed and resolved any uncertainties with the wider team. With a large author team potentially introducing heterogeneity in the process, we set up smaller groups of review authors working on each specific task (e.g. screening, extraction) to minimise inconsistencies. We also organised calls and discussions within these groups, where needed, to discuss any issues and harmonise the review process.

AUTHORS' CONCLUSIONS

Implications for practice

This review suggests that travel-related control measures during the COVID-19 pandemic may have a positive impact on infectious disease transmission and screening-related outcomes. However, the certainty of evidence included in this rapid review is moderate to very low, due to the nature as well as quality and heterogeneity of available studies. Therefore, true effects may be substantially different from those reported here. Broadly, travel restrictions reducing or stopping cross-border travel may limit the spread of disease across national borders. Regarding screening at borders, symptom/exposure screening on its own will detect some COVID-19 cases. However, it would likely not detect a large enough proportion of cases to prevent seeding new cases within the region protected by the measure. In comparison, PCR testing as a screening measure detects more cases than symptom/exposure screening, although if performed only upon arrival will likely also miss a

significant proportion of cases. The effectiveness of quarantine is dependent on high compliance and the length of quarantine, with longer periods such as 10 or 14 days preventing most cases from being released into the community. Combining quarantine with screening at borders is likely to meaningfully improve the effectiveness.

Travel-related control measures target one specific source of SARS-CoV-2 transmission, that is, human travel. The importance of this source of transmission in influencing overall epidemic development depends on a variety of factors, including the degree of interconnectedness between countries (in terms of the number and nature of borders as well as travel volumes) and the levels of community transmission in the region restricted and the region protected by the measure (i.e. high versus low levels of community transmission). Similarly, the contribution of travel-related control measures to controlling the COVID-19 pandemic will depend on the specification of these measures regarding their design and stringency (e.g. single polymerase chain reaction (PCR) testing versus repeated PCR testing; 7-day quarantine versus 14-day quarantine), the target group (e.g. all travellers versus specific groups), which borders or means of transport are affected (e.g. travel by air, land or sea), timing of implementation (measure implemented at an early versus late stage of the epidemic), and combinations with measures during travel (e.g. wearing of masks, hygiene, physical distancing). Importantly, the degree of adherence and, where applied, enforcement of the measure (e.g. recommendation to quarantine, various forms of control, fine and magnitude of fine) are also likely to play an important role. Finally, the contribution of travel-related control measures to controlling the COVID-19 pandemic will also depend on a range of other measures implemented to control community transmission (e.g. testing, contact tracing, social distancing measures, wearing face masks) in the region restricted and the region protected by the measure.

As the pandemic progresses, decision-makers can implement/increase or de-implement/loosen a range of potentially appropriate measures, and, in doing so, consider the above described factors. Importantly, travel-related control measures affect health and society in much broader ways, and decisions will need to balance all benefits and potential harms associated with a specific measure (not assessed in this review).

Implications for research

Decision-makers need high-quality research that helps to inform the decisions they continually have to make to chart the course through the COVID-19 pandemic. Research should be responsive to the questions most urgently raised by decision-makers, such as how can travel restrictions help delay a next wave of infections, how do screening at borders and quarantine need to be specified and enforced to optimise benefit-harm balance, and at what point is it safe to relax travel-related control measures. It would also be important that studies continue to refine the assessment of factors that influence the effectiveness of travel-related control measures, such as the stage of the pandemic and steps taken to increase or enforce implementation and adherence. Some of these questions may be answered quickly by refining existing models. Although measures such as border closures are, by nature, challenging if at all possible to evaluate using internally valid experimental or quasi-experimental approaches, the pandemic presents an opportunity to explore how critical questions can be answered through rigorous data collection and analysis.

In observational studies assessing the effectiveness of border screening measures and quarantine of travellers, it would be helpful to go beyond evaluating single measures. If, for example, all arriving travellers are quarantined for 14 days, it provides a cohort and an opportunity to assess not only the benefit of this quarantine, but also of a range of other single and combined measures, such as 3-day, 5-day and 8-day quarantine or PCR testing upon arrival, at day 3, day 5 or day 8. Additionally, related to outcomes it would be critical to look beyond the number of cases detected and to consider the number of cases missed, as well as the impact of these measures on the spread and development of the epidemic in the community. As many governments currently employ travel-related control measures, as well as a range of other public health measures to contain the pandemic, moving forward, it would be important to assess various combinations of these measures to identify those that are most effective. This concerns both primary research and future systematic reviews.

With the evidence base on COVID-19 and the impact of travel-related control measures growing quickly, it is important that future modelling studies improve reporting and technical documentation to allow for adequate assessment of their quality. Specific aspects considered in assessing the modelling studies in this review can help inform the development of high quality models. Finally, to ensure that the best available evidence informs complex and evolving decisions, future research should employ a range of epidemiological designs and assessment tools to assess the broad impacts of travel-related control measures, including all potential benefits and harms from a societal perspective. In order to integrate the rapidly growing evidence base on the topic, as well as new studies that may use more rigorous methods and approaches, we plan to update this review again later in 2021.

ACKNOWLEDGEMENTS

REFERENCES

References to studies included in this review

Adekunle 2020 {published data only}

Adekunle A, Meehan M, Rojas-Alvarez D, Trauer J, McBryde E. Delaying the COVID-19 epidemic in Australia: evaluating the effectiveness of international travel bans. *Australian and New Zealand Journal of Public Health* 2020;**44**(4):257-9. [DOI: 10.1111/1753-6405.13016]

Al-Qahtani 2020 {published data only}

Al-Qahtani M, AlAli S, Abdul RAK, Salman AA, Otoom S, Atkin SL. The prevalence of asymptomatic and symptomatic COVID19 disease in a cohort of quarantined subjects. International Journal of Infectious Diseases 2020 Nov 3 [Epub ahead of print].

Al-Tawfiq 2020 {published data only}

Al-Tawfiq JA, Sattar A, Al-Khadra H, Al-Qahtani S, Al-Mulhim M, Al-Omoush O, et al. Incidence of COVID-19 among returning travelers in quarantine facilities: A longitudinal study and lessons learned. *Travel Medicine and Infectious Disease* 2020;**38**:101901.

Anderson 2020 {published data only}

Anderson SC, Mulberry N, Edwards AM, Stockdale JE, Iyaniwura SA, Falcao RC, et al. How much leeway is there to relax COVID-19 control measures? *medRxiv* 2020. [DOI: 10.1101/2020.06.12.20129833]

Anzai 2020 {published data only}

Anzai A, Kobayashi T, Linton NM, Kinoshita R, Hayashi K, Suzuki A, et al. Assessing the impact of reduced travel on exportation dynamics of novel coronavirus infection (COVID-19). *Journal of Clinical Medicine* 2020;**9**(2):24. [DOI: dx.doi.org/10.3390/jcm9020601]

Arima 2020 {published data only}

* Arima Y, Shimada T, Suzuki M, Suzuki T, Kobayashi Y, Tsuchihashi Y, et al. Severe acute respiratory syndrome coronavirus 2 infection among returnees to Japan from Wuhan, China, 2020. *Emerging Infectious Diseases* 2020;**26**(7):1596-600. [DOI: 10.3201/eid2607.200994]

Nishiura H, Kobayashi T, Miyama T, Suzuki A, Jung SM, Hayashi K, et al. Estimation of the asymptomatic ratio of novel coronavirus infections (COVID-19). *International Journal of Infectious Diseases* 2020;**94**:154-5. [DOI: 10.1016/j.ijid.2020.03.020]

Ashcroft 2020 {published data only}

Ashcroft P, Lehtinen S, Angst DC, Low N, Bonhoeffer S. Quantifying the impact of quarantine duration on COVID-19 transmission. *medRxiv* 2020.

Banholzer 2020 {published data only}

Banholzer N, Van Weenen E, Kratzwald B, Seeliger A, Tschernutter D, Bottrighi P, et al. Estimating the impact of non-pharmaceutical interventions on documented infections with COVID-19: a cross-country analysis. *medRxiv* 2020. [DOI: 10.1101/2020.04.16.20062141]

Bays 2020 {published data only}

Bays D, Bennett E, Finnie T. Investigating the potential benefit that requiring travellers to self-isolate on arrival may have upon the reducing of case importations during international outbreaks of influenza, SARS, Ebola virus disease and COVID-19. *medRxiv* 2020.

* Bays D, Bennett E, Finnie T. Using simulation to assess the potential effectiveness of implementing screening at national borders during international outbreaks of influenza, SARS, Ebola virus disease and COVID-19. *medRxiv* 2020.

Binny 2020 {published data only}

Binny R, Baker MG, Hendy SC, James A, Lustig A, Plank MJ, et al. Early intervention is the key to success in COVID-19 control (preprint). *medRxiv* 2020.

Boldog 2020 {published data only}

Boldog P, Tekeli T, Vizi Z, Denes A, Bartha FA, Rost G. Risk assessment of novel coronavirus COVID-19 outbreaks outside China. *Journal of Clinical Medicine* 2020;**9**(2):19. [DOI: dx.doi.org/10.3390/jcm9020571]

Chen J 2020 {published data only}

Chen J, He H, Cheng W, Liu Y, Sun Z, Chai C, et al. Potential transmission of SARS-CoV-2 on a flight from Singapore to Hangzhou, China: An epidemiological investigation. *Travel Medicine and Infectious Disease* 2020;**36**:101816.

Chen T 2020 {published data only}

Chen T, Huang S, Li G, Zhang Y, Li Y, Zhu J, et al. Quantitative effects of entry restrictions and travel quarantine on the next wave of COVID-19: Case studies of China and Singapore. *SSRN* 2020.

Chen Y-H 2020 {published data only}

Chen Y-H, Fang C-T. Combined interventions to suppress R0 and border quarantine to contain COVID-19 in Taiwan. *Journal of the Formosan Medical Association* 2020;**120**(2):903-5.

Chinazzi 2020 {published data only}

Chinazzi M, Davis JT, Ajelli M, Gioannini C, Litvinova M, Merler S, et al. The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. *Science* 2020;**368**(6489):395-400.

Clifford 2020a {published data only}

Clifford S, Pearson CA, Klepac P, Van Zandvoort K, Quilty BJ, Eggo RM, et al, CMMID COVID-19 Working Group. Effectiveness of interventions targeting air travellers for delaying local outbreaks of SARS-CoV-2. *Journal of Travel Medicine* 2020;**8**:taaa068. [DOI: dx.doi.org/10.1093/jtm/taaa068]

Clifford 2020b {published data only}

Clifford S, Quilty BJ, Russell TW, Liu Y, Chan Yung-Wai D, Pearson CA, et al. Strategies to reduce the risk of SARS-CoV-2 re-introduction from international travellers. *medRxiv* 2020.

Costantino 2020 *{published data only}*

Costantino V, Heslop DJ, MacIntyre CR. The effectiveness of full and partial travel bans against COVID-19 spread in Australia for travellers from China during and after the epidemic peak in China. Journal of Travel Medicine 2020 May 22 [Epub ahead of print]. [DOI: dx.doi.org/10.1093/jtm/taaa081]

Davis 2020 *{published data only}*

Davis JT, Chinazzi M, Perra N, Mu K, Piontti PY, Ajelli M, et al. Estimating the establishment of local transmission and the cryptic phase of the COVID-19 pandemic in the USA. *medRxiv* 2020.

Deeb 2020 *{published data only}*

Deeb OE, Jalloul M. The dynamics of COVID-19 spread: evidence from Lebanon. *Mathematical Biosciences and Engineering* 2020;**17**(5):5618-32.

Dickens 2020 *{published data only}*

Dickens BL, Koo JR, Lim JT, Sun H, Clapham HE, Wilder-Smith A, et al. Strategies at points of entry to reduce importation risk of COVID-19 cases and re-open travel. *Journal of Travel Medicine* 2020;**27**(8):taaa141.

Gostic 2020 *{published data only}*

Gostic K, Gomez AC, Mummah RO, Kucharski AJ, Lloyd-Smith JO. Estimated effectiveness of symptom and risk screening to prevent the spread of COVID-19. *eLife* 2020;**0**(02):24. [DOI: dx.doi.org/10.7554/eLife.55570]

Grannell 2020 *{published data only}*

Grannell JJ, Grannell JR. A two-region SEIR COVID-19 epidemic model for the island of Ireland. *medRxiv* 2020.

Hoehl 2020 *{published data only}*

Hoehl S, Rabenau H, Berger A, Kortenbusch M, Cinatl J, Bojkova D, et al. Evidence of SARS-CoV-2 infection in returning travelers from Wuhan, China. *New England Journal of Medicine* 2020;**382**(13):1278-80. [DOI: dx.doi.org/10.1056/NEJMc2001899]

James 2020 *{published data only}*

James A, Plank MJ, Hendy S, Binny H, Lustig A, Steyn N. Model-free estimation of COVID-19 transmission dynamics from a complete outbreak. *medRxiv* 2020.

Kang 2020 *{published data only}*

Kang N, Kim B. The effects of border shutdowns on the spread of COVID-19. *Journal of Preventive Medicine and Public Health = Yebang Uihakhoe chi* 2020;**53**(5):293-301.

Kim 2020 *{published data only}*

Kim JG, Lee SH, Kim H, Oh HS, Lee J. Air evacuation of passengers with potential SARS-CoV-2 infection under the guidelines for appropriate infection control and prevention. *Osong Public Health and Research Perspectives* 2020;**11**(5):334-8.

Kivuti-Bito 2020 *{published data only}*

Kivuti-Bitok LW, Momodu AS, Cheptum JJ, Kimemia F, Gichuki I, Ngune I. System dynamics model of possible Covid-19

trajectories under various non-pharmaceutical intervention options in low-resource setting. *medRxiv* 2020.

Kwok 2020 *{published data only}*

Kwok WC, Wong CK, Ma TF, Ho KW, Fan WTL, Chan KPF, et al. Border Restriction as a Public Health Measure to Limit Outbreak of Coronavirus Disease 2019 (COVID-19). *medRxiv* 2020.

Lagier 2020 *{published data only}*

Lagier JC, Colson P, Tissot Dupont H, Salomon J, Doudier B, Aubry C, et al. Testing the repatriated for SARS-Cov2: Should laboratory-based quarantine replace traditional quarantine? *Travel Medicine and Infectious Disease* 2020;**34**:101624.

Liebig 2020 *{published data only}*

Liebig J, Najeebullah K, Jurdak R, Shoghri AE, Paini D. Should international borders re-open? The impact of travel restrictions on COVID-19 importation risk. *medRxiv* 2020.

Linka 2020a *{published data only}*

Linka K, Peirlinck M, Kuhl E. The reproduction number of COVID-19 and its correlation with public health interventions. *Computational Mechanics* 2020;**66**(4):1035-50.

* Linka K, Peirlinck M, Sahli Costabal F, Kuhl E. Outbreak dynamics of COVID-19 in Europe and the effect of travel restrictions. *Computer Methods in Biomechanics & Biomedical Engineering, 1-8* 2020;**23**(11):710-17. [DOI: 10.1080/10255842.2020.1759560]

Linka 2020b *{published data only}*

Linka K, Rahman P, Goriely A, Kuhl E. Is it safe to lift COVID-19 travel bans? The Newfoundland story. Computational Mechanics 2020 Aug 29 [Epub ahead of print].

Lio 2020 *{published data only}*

Lio CF, Cheong HH, Lei CI, Lo IL, Yao L, Lam C, et al. The common personal behavior and preventive measures among 42 uninfected travelers from the Hubei province, China during COVID-19 outbreak: a cross-sectional survey in Macao SAR, China. *PeerJ* 2020;**8**:e9428.

Lytras 2020 *{published data only}*

Lytras T, Dellis G, Flountzi A, Hatzianastasiou S, Nikolopoulou G, Tsekou K, et al. High prevalence of SARS-CoV-2 infection in repatriation flights to Greece from three European countries. *Journal of Travel Medicine* 2020;**27**(3):18. [DOI: dx.doi.org/10.1093/jtm/taaa054]

Mandal 2020 *{published data only}*

Mandal S, Bhatnagar T, Arinaminpathy N, Agarwal A, Chowdhury A, Murhekar M, et al. Prudent public health intervention strategies to control the coronavirus disease 2019 transmission in India: a mathematical model-based approach. *Indian Journal of Medical Research* 2020;**151**(2 & 3):190-9. [DOI: dx.doi.org/10.4103/ijmr.IJMR_504_20]

McLure 2020 *{published data only}*

McLure A, Lau LL, Furuya-Kanamori L. Has the effectiveness of Australia's travel bans against China on the importation of

COVID-19 been overestimated? Journal of Travel Medicine 2020 Oct 10 [Epub ahead of print].

Nakamura 2020 *{published data only}*

Nakamura H, Managi S. Airport risk of importation and exportation of the COVID-19 pandemic. *Transport Policy* 2020;**96**:40-7.

Ng 2020 *{published data only}*

Ng OT, Marimuthu K, Chia PY, Koh V, Chiew CJ, De Wang L, et al. SARS-CoV-2 infection among travelers returning from Wuhan, China. *New England Journal of Medicine* 2020;**382**(14):1476-8. [DOI: dx.doi.org/10.1056/NEJMc2003100]

Nowrasteh 2020 *{published data only}*

Nowrasteh A, Forrester AC, Cato Institute. How US travel restrictions on China affected the spread of COVID-19 in the United States. www.cato.org 2020.

Nuckchady 2020 *{published data only}*

Nuckchady D. Impact of public health interventions on the COVID-19 epidemic: a stochastic model based on data from an African island. *medRxiv* 2020.

Odendaal 2020 *{published data only}*

Odendaal WG. A method to model outbreaks of new infectious diseases with pandemic potential such as COVID-19. *medRxiv* 2020. [DOI: 10.1101/2020.03.11.20034512]

Pinotti 2020 *{published data only}*

Pinotti F, Di Domenico L, Ortega E, Mancastroppa M, Pullano G, Valdano E, et al. Tracing and analysis of 288 early SARS-CoV-2 infections outside China: A modeling study. *PLoS medicine* 2020;**17**(7):e1003193.

Quilty 2020 *{published data only}*

Quilty BJ, Clifford S, Flasche S, Eggo RM, CCMID nCoV Working Group. Effectiveness of airport screening at detecting travellers infected with novel coronavirus (2019-nCoV). *Eurosurveillance* 2020;**25**(5):2. [DOI: dx.doi.org/10.2807/1560-7917.ES.2020.25.5.2000080]

Russell TW 2020 *{published data only}*

Russell TW, Wu JT, Clifford S, Edmunds WJ, Kucharski AJ, Jit M. Effect of internationally imported cases on internal spread of COVID-19: a mathematical modelling study. *Lancet Public Health* 2021;**6**(1):e12-20.

Russell WA 2020 *{published data only}*

Russell WA, Buckeridge DL. Effectiveness of quarantine and testing to prevent COVID-19 transmission from arriving travelers. *medRxiv* 2020.

Ryu 2020 *{published data only}*

Ryu S, Ali ST, Lim JS, Chun BC. Estimation of the excess COVID-19 cases in Seoul, South Korea by the students arriving from China. *International Journal of Environmental Research and Public Health* 2020;**17**:3113. [DOI: dx.doi.org/10.3390/ijerph17093113]

Shaikh Abdul Karim 2020 *{published data only}*

Shaikh Abdul Karim S, Md Tahir FA, Mohamad UK, Abu Bakar M, Mohamad KN, Suleiman M, et al. Experience repatriation of citizens from epicentre using commercial flights during COVID-19 pandemic. *International Journal of Emergency Medicine* 2020;**13**(1):50.

Shi 2020 *{published data only}*

Shi S, Tanaka S, Ueno R, Gilmour S, Tanoue Y, Kawashima T, et al. Travel restrictions and SARS-CoV-2 transmission: an effective distance approach to estimate impact. *Bulletin of the World Health Organization* 2020;**98**(8):518-29.

Sruthi 2020 *{published data only}*

Sruthi CK, Biswal MR, Saraswat B, Joshi H, Prakash MK. How policies on restaurants, bars, nightclubs, masks, schools, and travel influenced Swiss COVID-19 reproduction ratios. *medRxiv* 2020.

Steyn 2020 *{published data only}*

Steyn N, Plank MJ, James A, Binny RN, Hendy S, Lustig A. Managing the risk of a COVID-19 outbreak from border arrivals. *medRxiv* 2020.

Taylor 2020 *{published data only}*

Taylor R, McCarthy CA, Patel V, Moir R, Kelly L, Snary E. The risk of introducing SARS-CoV-2 to the UK via international travel in August 2020. *medRxiv* 2020.

Utsunomiya 2020 *{published data only}*

Utsunomiya YT, Utsunomiya AT, Torrecilha RB, Paulan SC, Milanesi M, Garcia JF. Growth rate and acceleration analysis of the COVID-19 pandemic reveals the effect of public health measures in real time. *Frontiers in Medicine* 2020;**7**:247. [DOI: 10.3389/fmed.2020.00247]

Wells 2020 *{published data only}*

Wells CR, Sah P, Moghadas SM, Pandey A, Shoukat A, Wang Y, et al. Impact of international travel and border control measures on the global spread of the novel 2019 coronavirus outbreak. *Proceedings of the National Academy of Sciences of the United States of America* 2020;**117**(13):7504-9. [DOI: dx.doi.org/10.1073/pnas.2002616117]

Wilson 2020 *{published data only}*

Wilson N, Baker MG, Eichner M. Estimating the impact of control measures to prevent outbreaks of COVID-19 associated with air travel into a COVID-19-free country: a simulation modelling study. *medRxiv* 2020. [DOI: 10.1101/2020.06.10.20127977]

Wong J 2020 *{published data only}*

Wong J, Abdul Aziz ABZ, Chaw L, Mahamud A, Griffith MM, Lo YR, et al. High proportion of asymptomatic and presymptomatic COVID-19 infections in air passengers to Brunei. *Journal of Travel Medicine* 2020;**27**(5):taaa066.

Wong MC 2020 *{published data only}*

Wong MC, Ng RW, Chong KC, Lai CK, Huang J, Chen Z, et al. Stringent containment measures without complete city lockdown to achieve low incidence and mortality across two waves of COVID-19 in Hong Kong. BMJ Global Health 2020;**5**(10).

Cochrane
Library

Trusted evidence.
Informed decisions.
Better health.

Cochrane Database of Systematic Reviews

Yamahata 2020 *{published data only}*

Tsuboi M, Hachiya M, Noda S, Iso H, Umeda T. Epidemiology and quarantine measures during COVID-19 outbreak on the cruise ship Diamond Princess docked at Yokohama, Japan in 2020: a descriptive analysis. *Global Health & Medicine* 2020;**2**(2):102-6. [DOI: 10.35772/ghm.2020.01037]

Yamahata Y, Shibata A. Preparation for quarantine on the cruise ship Diamond Princess in Japan due to COVID-19. *JMIR Public Health and Surveillance* 2020;**6**(2):e18821. [DOI: dx.doi.org/10.2196/18821]

Yang 2020 *{published data only}*

Yang T, Liu Y, Deng W, Zhao W, Deng J. SARS-Cov-2 trajectory predictions and scenario simulations from a global perspective: a modelling study. *Scientific Reports* 2020;**10**(1):18319.

Zhang C 2020 *{published data only}*

Zhang C, Qian LX, Hu JQ. COVID-19 pandemic with human mobility across countries. Journal of the Operations Research Society of China 2020 Aug 3 [Epub ahead of print].

Zhang L 2020 *{published data only}*

Zhang L, Yang H, Wang K, Zhan Y, Bian L. Measuring imported case risk of COVID-19 from inbound international flights - A case study on China. *Journal of Air Transport Management* 2020;**89**:101918.

Zhong 2020 *{published data only}*

Zhong L, Diagne M, Wang W, Gao J. Country distancing reveals the effectiveness of travel restrictions during COVID-19. *medRxiv* 2020.

References to studies excluded from this review

Adiga 2020 *{published data only}*

Adiga A, Wang L, Sadilek A, Tendulkar A, Venkatramanan S, Vullikanti A, et al. Interplay of global multi-scale human mobility, social distancing, government interventions, and COVID-19 dynamics. *medRxiv* 2020. [DOI: dx.doi.org/10.1101/2020.06.05.20123760]

Aleta 2020 *{published data only}*

Aleta A, Hu Q, Ye J, Ji P, Moreno Y. A data-driven assessment of early travel restrictions related to the spreading of the novel COVID-19 within mainland China. *Chaos, Solitons & Fractals* 2020;**139**:110068.

Annan 2015 *{published data only}*

Annan A, Owusu M, Marfo KS, Larbi R, Sarpong FN, Adu-Sarkodie Y, et al. High prevalence of common respiratory viruses and no evidence of Middle East respiratory syndrome coronavirus in Hajj pilgrims returning to Ghana, 2013. *Tropical Medicine & International Health* 2015;**20**(6):807-12.

Aravindakshan 2020 *{published data only}*

Aravindakshan A, Boehnke J, Gholami E, Nayak A. Restarting after COVID-19: A data-driven evaluation of opening scenarios. *medRxiv* 2020. [DOI: 10.1101/2020.05.28.20115980]

Arino 2020 *{published data only}*

Arino J, Bajeux N, Portet S, Watmough J. Assessing the risk of COVID-19 importation and the effect of quarantine. *medRxiv* 2020.

Barkan 2020 *{published data only}*

Barkan E, Shilo S, Talmor-Barkan Y. Comparison of SARS-CoV-2 exit strategies building blocks. *medRxiv* 2020. [DOI: 10.1101/2020.04.23.20072850]

Batista 2020 *{published data only}*

Batista B, Dickenson D, Gurski K, Kebe M, Rankin N. Minimizing disease spread on a quarantined cruise ship: A model of COVID-19 with asymptomatic infections. *Mathematical Biosciences* 2020;**329**:108442.

Bell 2004 *{published data only}*

Bell DM, World Health Organization Working Group on International Community Transmission of SARS . Public health interventions and SARS spread, 2003. *Emerging Infectious Diseases* 2004;**10**(11):1900-6.

Benkouiten 2013 *{published data only}*

Benkouiten S, Charrel R, Belhouchat K, Drali T, Salez N, Nougairede A, et al. Circulation of respiratory viruses among pilgrims during the 2012 Hajj pilgrimage. *Clinical Infectious Diseases* 2013;**57**(7):992-1000.

Benkouiten 2014 *{published data only}*

Benkouiten S, Charrel R, Belhouchat K, Drali T, Nougairede A, Salez N, et al. Respiratory viruses and bacteria among pilgrims during the 2013 Hajj. *Emerging Infectious Diseases* 2014;**20**(11):1821-7.

Benkouiten 2015 *{published data only}*

Benkouiten S, Gautret P, Belhouchat K, Drali T, Nougairede A, Salez N, et al. Comparison of nasal swabs with throat swabs for the detection of respiratory viruses by real-time reverse transcriptase PCR in adult Hajj pilgrims. *Journal of Infection* 2015;**70**(2):207-10.

Borracci 2020 *{published data only}*

Borracci RA, Giglio ND. Forecasting the effect of social distancing on COVID-19 autumn-winter outbreak in the metropolitan area of Buenos Aires. *Estimacion del efecto del distanciamiento social sobre la epidemia de COVID-19 de otono-invierno en el area metropolitana de Buenos Aires.* 2020;**80 Suppl 3**:7-15.

Branas 2020 *{published data only}*

Branas CC, Rundle A, Pei S, Yang W, Carr BG, Sims S, et al. Flattening the curve before it flattens us: hospital critical care capacity limits and mortality from novel coronavirus (SARS-CoV2) cases in US counties. *medRxiv* 2020. [DOI: 10.1101/2020.04.01.20049759]

Brauer 2008 *{published data only}*

Brauer F, Van den Driessche P, Wang L. Oscillations in a patchy environment disease model. *Mathematical Biosciences* 2008;**215**(1):1-10.

Camitz 2006 {published data only}

Camitz M, Liljeros F. The effect of travel restrictions on the spread of a moderately contagious disease. *BMC Medicine* 2006;**4**:32.

Channapathi 2020 {published data only}

Channapathi T, Thatikonda S. Stochastic transmission dynamic model for evaluating effectiveness of control measures of COVID-19. *SSRN* 2020. [DOI: dx.doi.org/10.2139/ssrn.3559581]

Cowling 2020 {published data only}

Cowling BJ, Ali ST, Ng TW, Tsang TK, Li JC, Fong MW, et al. Impact assessment of non-pharmaceutical interventions against coronavirus disease 2019 and influenza in Hong Kong: an observational study. *Lancet Public Health* 2020;**5**(5):e279-88.

Dandekar 2020 {published data only}

Dandekar R, Barbastathis G. Quantifying the effect of quarantine control in COVID-19 infectious spread using machine learning. *medRxiv* 2020. [DOI: 10.1101/2020.04.03.20052084]

Daon 2020 {published data only}

Daon Y, Thompson RN, Obolski U. Estimating COVID-19 outbreak risk through air travel. *medRxiv* 2020.

Dell'Omodarme 2005 {published data only}

Dell'Omodarme M, Prati MC. The probability of failing in detecting an infectious disease at entry points into a country. *Statistics in Medicine* 2005;**24**(17):2669-79.

Dursun 2020 {published data only}

Dursun ZB, Ulu-Kilic A, Alabay S, Benli AR, Celik I. COVID-19 among Turkish citizens returning from abroad. *Travel Medicine and Infectious Disease* 2020;**37**:101860.

Eksin 2020 {published data only}

Eksin C, Ndeffo-Mbah M, Weitz JS. Reacting to outbreaks at neighboring localities. *medRxiv* 2020. [DOI: 10.1101/2020.04.24.20078808]

Erandi 2020 {published data only}

Erandi KK, Mahasinghe AC, Perera SS, Jayasinghe S. Effectiveness of the strategies implemented in Sri Lanka for controlling the COVID-19 outbreak. *Journal of Applied Mathematics* 2020;**2020**:2954519.

Espinoza 2020 {published data only}

Espinoza B, Castillo-Chavez C, Perrings C. Mobility restrictions for the control of epidemics: when do they work? *SSRN* 2020. [DOI: 10.2139/ssrn.3496928]

Fang 2020 {published data only}

Fang Y, Nie Y, Penny M. Transmission dynamics of the COVID-19 outbreak and effectiveness of government interventions: a data-driven analysis. *Journal of Medical Virology* 2020. [DOI: 10.1002/jmv.25750]

Fouquet 2020 {published data only}

Fouquet R, O'Garra T. The behavioural, welfare and environmental effects of air travel reductions during and beyond COVID-19. *SSRN* 2020. [DOI: dx.doi.org/10.2139/ssrn.3628750]

Fredj 2020 {published data only}

Fredj HB, Cherif F. Novel corona virus disease infection in Tunisia: mathematical model and the impact of the quarantine strategy. *Chaos Solitons & Fractals* 2020;**138**:109969.

Gardner 2016 {published data only}

Gardner LM, Chughtai AA, MacIntyre CR. Risk of global spread of Middle East respiratory syndrome coronavirus (MERS-CoV) via the air transport network. *Journal of Travel Medicine* 2016;**23**(6):taw063. [DOI: 10.1093/jtm/taw063]

Gatto 2020 {published data only}

Gatto M, Bertuzzo E, Mari L, Miccoli S, Carraro L, Casagrandi R, et al. Spread and dynamics of the COVID-19 epidemic in Italy: effects of emergency containment measures. *Proceedings of the National Academy of Sciences of the United States of America* 2020. [DOI: 10.1073/pnas.2004978117]

Gautret 2013a {published data only}

Gautret P, Benkouiten S, Salaheddine I, Parola P, Brouqui P. Preventive measures against MERS-CoV for Hajj pilgrims. *Lancet Infectious Diseases* 2013;**13**(10):829-31.

Gautret 2013b {published data only}

Gautret P, Charrel R, Belhouchat K, Drali T, Benkouiten S, Nougairede A, et al. Lack of nasal carriage of novel corona virus (HCoV-EMC) in French Hajj pilgrims returning from the Hajj 2012, despite a high rate of respiratory symptoms. *Clinical Microbiology & Infection* 2013;**19**(7):E315-7.

Gautret 2014 {published data only}

Gautret P, Charrel R, Benkouiten S, Belhouchat K, Nougairede A, Drali T, et al. Lack of MERS coronavirus but prevalence of influenza virus in French pilgrims after 2013 Hajj. *Emerging Infectious Diseases* 2014;**20**(4):728-30.

German 2015 {published data only}

German M, Olsha R, Kristjanson E, Marchand-Austin A, Peci A, Winter AL, et al. Acute respiratory infections in travelers returning from MERS-CoV-affected areas. *Emerging Infectious Diseases* 2015;**21**(9):1654-6.

Gill 2020 {published data only}

Gill BS, Jayaraj VJ, Singh S, Ghazali MS, Cheong YL, Md Iderus NH, et al. Modelling the effectiveness of epidemic control measures in preventing the transmission of COVID-19 in Malaysia. *International Journal of Environmental Research and Public Health* 2020;**17**(15):5509.

Godin 2021 {published data only}

Godin A, Xia Y, Buckeridge DL, Mishra S, Douwes-Schultz D, Shen Y, et al. The role of case importation in explaining differences in early SARS-CoV-2 transmission dynamics in Canada - a mathematical modeling study of surveillance data. *International Journal of Infectious Diseases* 2021;**102**:254-9.

Griffiths 2016 *{published data only}*

Griffiths K, Charrel R, Lagier JC, Nougairede A, Simon F, Parola P, et al. Infections in symptomatic travelers returning from the Arabian peninsula to France: a retrospective cross-sectional study. *Travel Medicine & Infectious Disease* 2016;**14**(4):414-6.

Gunthe 2020 *{published data only}*

Gunthe SS, Patra SS. Impact of international travel dynamics on domestic spread of 2019-nCoV in India: origin-based risk assessment in importation of infected travelers. *Global Health* 2020;**16**(1):45.

Hossain 2020 *{published data only}*

Hossain MP, Junus A, Zhu X, Jia P, Wen TH, Pfeiffer D, et al. The effects of border control and quarantine measures on the spread of COVID-19. *Epidemics* 2020;**32**:100397.

Hossein 2020 *{published data only}*

Hossein RT, Shahriari S, Azad AK, Vafaee F. Real-time time-series modelling for prediction of COVID-19 spread and intervention assessment. *medRxiv* 2020. [DOI: 10.1101/2020.04.24.20078923]

Huang 2020 *{published data only}*

Huang L, Li L, Dunn L, He M. Taking account of asymptomatic infections in modeling the transmission potential of the COVID-19 outbreak on the Diamond Princess cruise ship. *medRxiv* 2020.

Hufnagel 2004 *{published data only}*

Hufnagel L, Brockmann D, Geisel T. Forecast and control of epidemics in a globalized world. *Proceedings of the National Academy of Sciences of the United States of America* 2004;**101**(42):15124-9.

Jia 2020 *{published data only}*

Jia JS, Lu X, Yuan Y, Xu G, Jia J, Christakis NA. Population flow drives spatio-temporal distribution of COVID-19 in China. *Nature* 2020;**582**(7812):389-94.

Johansson 2011 *{published data only}*

Johansson MA, Arana-Vizcarrondo N, Biggerstaff BJ, Staples JE, Gallagher N, Marano N. On the treatment of airline travelers in mathematical models. *PLoS ONE* 2011;**6**(7):e22151.

Joo 2019 *{published data only}*

Joo H, Maskery BA, Berro AD, Rotz LD, Lee YK, Brown CM. Economic impact of the 2015 MERS outbreak on the Republic of Korea's tourism-related industries. *Health Security* 2019;**17**(2):100-8.

Jungerman 2017 *{published data only}*

Jungerman MR, Vonnahme LA, Washburn F, Alvarado-Ramy F. Federal travel restrictions to prevent disease transmission in the United States: an analysis of requested travel restrictions. *Travel Medicine & Infectious Disease* 2017;**18**:30-5.

Kong 2020 *{published data only}*

Kong XS, Liu F, Wang HB, Yang RF, Chen DB, Wang XX, et al. Epidemic prevention and control measures in China significantly curbed the epidemic of COVID-19 and influenza. *medRxiv* 2020. [DOI: 10.1101/2020.04.09.20058859]

Kraemer 2020 *{published data only}*

Kraemer MU, Yang CH, Gutierrez B, Wu CH, Klein B, Pigott DM, et al. The effect of human mobility and control measures on the COVID-19 epidemic in China. *Science* 2020;**368**(6490):493-7.

Krisztin 2020 *{published data only}*

Krisztin T, Piribauer P, Woegerer M. The spatial econometrics of the coronavirus pandemic. *Letters in Spatial and Resource Sciences* 2020;**13**(3):209-18.

Lai CC 2020 *{published data only}*

Lai CC, Hsu CY, Jen HH, Yen MF, Chan CC, Chen HH. Bayesian approach for modelling the dynamic of COVID-19 outbreak on the Diamond Princess Cruise Ship. *medRxiv* 2020.

Lai S 2020a *{published data only}*

Lai S, Bogoch I, Ruktanonchai N, Watts A, Lu X, Yang W, et al. Assessing spread risk of Wuhan novel coronavirus within and beyond China, January-April 2020: a travel network-based modelling study. *medRxiv* 2020. [DOI: dx.doi.org/10.1101/2020.02.04.20020479]

Lai S 2020b *{published data only}*

Lai S, Ruktanonchai NW, Carioli A, Ruktanonchai C, Floyd J, Prosper O, et al. Assessing the effect of global travel and contact reductions to mitigate the COVID-19 pandemic and resurgence. *medRxiv* 2020. [DOI: 10.1101/2020.06.17.20133843]

Lai S 2020c *{published data only}*

Lai S, Ruktanonchai NW, Zhou L, Prosper O, Luo W, Floyd JR, et al. Effect of non-pharmaceutical interventions to contain COVID-19 in China. *Nature* 2020;**04**:04. [DOI: dx.doi.org/10.1038/s41586-020-2293-x]

Lam 2020 *{published data only}*

Lam HY, Lam TS, Wong CH, Lam WH, Leung CM, Au KW, et al. The epidemiology of COVID-19 cases and the successful containment strategy in Hong Kong - January to May 2020. *International Journal of Infectious Diseases* 2020;**21**:21. [DOI: dx.doi.org/10.1016/j.ijid.2020.06.057]

Lau 2004 *{published data only}*

Lau JT, Yang X, Tsui HY, Pang E. SARS related preventive and risk behaviours practised by Hong Kong-mainland China cross border travellers during the outbreak of the SARS epidemic in Hong Kong. *Journal of Epidemiology & Community Health* 2004;**58**(12):988-96.

Lee 2020 *{published data only}*

Lee VJ, Chiew CJ, Khong WX. Interrupting transmission of COVID-19: lessons from containment efforts in Singapore. *Journal of Travel Medicine* 2020;**27**(3):18.

Li H 2020 *{published data only}*

Li H, Sun K, Persing DH, Tang YW, Shen D. Real-time screening of specimen pools for coronavirus disease 2019 (COVID-19) infection at Sanya Airport, Hainan Island, China. *Clinical Infectious Diseases* 2020 Sep 30 [Epub ahead of print].

Lin YC 2020 *{published data only}*

Lin YC, Chen MY, Liu MC, Lin YJ, Lin Yu-H, Kuo JS, et al. Quarantine measures for coronavirus disease 2019 on a cruise ship, Taiwan, February 2020. *International Journal of Infectious Diseases* 2020;**99**:298-300.

Lin YH 2020 *{published data only}*

Lin YH, Huang JY, Yu KD, Lu CM, Lee WP, Huang JJ, et al. Border quarantine measures and achievement of COVID-19 control in Taiwan. *Epidemiology Bulletin* 2020;**36**(15):87-8.

Li R 2020 *{published data only}*

Li R, Chen B, Zhang T, Ren Z, Song Y, Xiao Y, et al. Global COVID-19 pandemic demands joint interventions for the suppression of future waves. *Proceedings of the National Academy of Sciences of the United States of America* 2020;**117**(42):26151-7.

Liu 2011 *{published data only}*

Liu X, Chen X, Takeuchi Y. Dynamics of an SIQS epidemic model with transport-related infection and exit-entry screenings. *Journal of Theoretical Biology* 2011;**285**(1):25-35.

Liu F 2020 *{published data only}*

Liu F, Wang M, Zheng M. Effects of COVID-19 lockdown on global air quality and health. *The Science of the Total Environment* 2020;**755**(Pt 1):142533.

Liu Q 2020 *{published data only}*

Liu Q, Lu H, Chen R. Effect of a bundle of intervention strategies for the control of COVID-19 in Henan, a neighboring province of Wuhan, China. *Wiener Klinische Wochenschrift* 2020;**132**(13-14):396-399. [DOI: 10.1007/s00508-020-01688-9]

Luna 2007 *{published data only}*

Luna LK, Panning M, Grywna K, Pfefferle S, Drosten C. Spectrum of viruses and atypical bacteria in intercontinental air travelers with symptoms of acute respiratory infection. *Journal of Infectious Diseases* 2007;**195**(5):675-9.

Ma 2017 *{published data only}*

Ma X, Liu F, Liu L, Zhang L, Lu M, Abudukadeer A, et al. No MERS-CoV but positive influenza viruses in returning Hajj pilgrims, China, 2013-2015. *BMC Infectious Diseases* 2017;**17**:715.

Maeno 2016 *{published data only}*

Maeno Y. Detecting a trend change in cross-border epidemic transmission. *Physica A* 2016;**457**:73-81.

Magalis 2020 *{published data only}*

Magalis BR, Ramirez-Mata A, Zhukova A, Mavian C, Marini S, Lemoine F, et al. Differing impacts of global and regional responses on SARS-CoV-2 transmission cluster dynamics. *bioRxiv* 2020.

Malmberg 2020 *{published data only}*

Malmberg H, Britton T. Inflow restrictions can prevent epidemics when contact tracing efforts are effective but have limited capacity. *Journal of The Royal Society Interface* 2020;**17**(170):20200351.

Marcelino 2012 *{published data only}*

Marcelino J, Kaiser M. Critical paths in a metapopulation model of H1N1: efficiently delaying influenza spreading through flight cancellation. *PLoS Currents* 2012;**4**:e4f8c9a2e1fca8.

Myers 2020 *{published data only}*

Myers JF, Snyder RE, Porse CC, Tecle S, Lowenthal P, Danforth ME, et al. Identification and monitoring of international travelers during the initial phase of an outbreak of COVID-19 - California, February 3-March 17, 2020. *MMWR - Morbidity & Mortality Weekly Report* 2020;**69**(19):599-602.

Ng 2020a *{published data only}*

Ng TCV, Cheng HY, Chang HH, Liu CC, Yang CC, Jian SW, et al. Effects of case- and population-based COVID-19 interventions in Taiwan. *medRxiv* 2020.

Nikolaou 2020 *{published data only}*

Nikolaou P, Dimitriou L. Identification of critical airports for controlling global infectious disease outbreaks: stress-tests focusing in Europe. *Journal of Air Transport Management* 2020;**85**:101819.

Niwa 2020 *{published data only}*

Niwa M, Hara Y, Sengoku S, Kodama K. Effectiveness of social measures against COVID-19 outbreaks in selected Japanese regions analyzed by system dynamic modeling. *International Journal of Environmental Research and Public Health* 2020;**17**(17):6238.

Pan 2020 *{published data only}*

Pan W, Tyrovolas S, Vazquez IG, Raj R, Fernandez D, Zaitchik B, et al. COVID-19: Effectiveness of non-pharmaceutical interventions in the United States before phased removal of social distancing protections varies by region. *medRxiv* 2020.

Pitman 2005 *{published data only}*

Pitman RJ, Cooper BS, Trotter CL, Gay NJ, Edmunds WJ. Entry screening for severe acute respiratory syndrome (SARS) or influenza: policy evaluation. *BMJ* 2005;**331**(7527):1242-3.

Pullano 2020 *{published data only}*

Pullano G, Valdano E, Scarpa N, Rubrichi S, Colizza V. Population mobility reductions during COVID-19 epidemic in France under lockdown. *medRxiv* 2020. [DOI: 10.1101/2020.05.29.20097097]

Quilty 2020b *{published data only}*

Quilty BJ, Diamond C, Liu Y, Gibbs H, Russell TW, Jarvis CI, et al. The effect of inter-city travel restrictions on geographical spread of COVID-19: evidence from Wuhan, China. *medRxiv* 2020. [DOI: 10.1101/2020.04.16.20067504]

Rajabi 2020 *{published data only}*

Rajabi A, Mantzaris AV, Mutlu EC, Garibay I. Investigating dynamics of COVID-19 spread and containment with agent-based modeling. *medRxiv* 2020.

Ruktanonchai 2020 *{published data only}*

Ruktanonchai NW, Floyd JR, Lai S, Ruktanonchai CW, Sadilek A, Rente-Lourenco P, et al. Assessing the impact of

coordinated COVID-19 exit strategies across Europe. *Science* 2020;**369**(6510):1465-70.

Ryu 2019 *{published data only}*

Ryu S, Kim JJ, Cowling BJ, Kim C. Surveillance and public health response for travelers returning from MERS-CoV affected countries to Gyeonggi Province, Korea, 2016-2017. *Travel Medicine & Infectious Disease* 2019;**31**:101350.

Shah 2020 *{published data only}*

Shah NH, Sheoran N, Jayswal EN, Shukla D, Shukla N, Shukla J, et al. Modelling COVID-19 transmission in the United States through interstate and foreign travels and evaluating impact of governmental public health interventions. *medRxiv* 2020. [DOI: 10.1101/2020.05.23.20110999]

Shumway 2020 *{published data only}*

Shumway B, Ibrahim D, Moss W. Monitoring returning travelers during the early weeks of the COVID-19 pandemic: one US county's experience. *American Journal of Public Health* 2020;**110**(7):962-3.

Sriwijitalai 2020b *{published data only}*

Sriwijitalai W, Wiwanitkit V. Positive screening for Wuhan novel coronavirus infection at international airport: what's the final diagnosis for positive cases. *International Journal of Preventive Medicine* 2020;**11**:30.

Summan 2020 *{published data only}*

Summan A, Nandi A. Timing of non-pharmaceutical interventions to mitigate COVID-19 transmission and their effects on mobility: a cross-country analysis. *medRxiv* 2020. [DOI: 10.1101/2020.05.09.20096420]

Sun 2013 *{published data only}*

Sun G, Abe N, Sugiyama Y, Nguyen QV, Nozaki K, Nakayama Y, et al. Development of an infection screening system for entry inspection at airport quarantine stations using ear temperature, heart and respiration rates. *Conference Proceedings: Annual International Conference of the IEEE Engineering in Medicine & Biology Society* 2013;**2013**:6716-9.

Thomas 2014 *{published data only}*

Thomas HL, Zhao H, Green HK, Boddington NL, Carvalho CF, Osman HK, et al. Enhanced MERS coronavirus surveillance of travelers from the Middle East to England. *Emerging Infectious Diseases* 2014;**20**(9):1562-4.

Valba 2020 *{published data only}*

Valba O, Avetisov V, Gorsky A, Nechaev S. Self-isolation or borders closing: What prevents the spread of the epidemic better? *Physical Review E* 2020;**102**(1-1):010401.

Wells 2020b *{published data only}*

Wells CR, Townsend JP, Pandey A, Krieger G, Singer BH, McDonald RH, et al. Optimal COVID-19 quarantine and testing strategies. *medRxiv* 2020.

Wickramaarachchi 2020 *{published data only}*

Wickramaarachchi WP, Perera SS, Jayasinghe S. COVID-19 epidemic in Sri Lanka: a mathematical and computational modelling approach to control. *Computational and Mathematical Methods in Medicine* 2020;**2020**:4045064.

Yip 2007 *{published data only}*

Yip PS, Hsieh YH, Xu Y, Lam KF, King CC, Chang HL. Assessment of intervention measures for the 2003 SARS epidemic in Taiwan by use of a back-projection method. *Infection Control and Hospital Epidemiology* 2007;**28**(5):525-30.

Yuan 2020 *{published data only}*

Yuan HY, Hossain MP, Tsegaye MM, Zhu X, Jia P, Wen TH, et al. Estimating the risk on outbreak spreading of 2019-nCoV in China using transportation data. *medRxiv* 2020. [DOI: 10.1101/2020.02.01.20019984]

Zhao Q 2020 *{published data only}*

Zhao Q, Chen Y, Small DS. Analysis of the epidemic growth of the early 2019-nCoV outbreak using internationally confirmed cases. *medRxiv* 2020. [DOI: 10.1101/2020.02.06.20020941]

Zhao Z 2020 *{published data only}*

Zhao Z, Li X, Liu F, Zhu G, Ma C, Wang L. Prediction of the COVID-19 spread in African countries and implications for prevention and control: a case study in South Africa, Egypt, Algeria, Nigeria, Senegal and Kenya. *Science of the Total Environment* 2020;**729**:138959.

Zheng 2020 *{published data only}*

Zheng Y. Estimation of disease transmission in multimodal transportation networks. Journal of Advanced Transportation 2020;**2020**. [DOI: 10.1155/2020/8898923]

Additional references

Arshed 2020

Arshed N, Meo MS, Farooq F. Empirical assessment of government policies and flattening of the COVID19 curve. *Journal of Public Affairs* 2020:e2333.

Baba 2020

Baba IA, Baba BA, Esmaili P. A mathematical model to study the effectiveness of some of the strategies adopted in curtailing the spread of COVID-19. *Computational and Mathematical Methods in Medicine* 2020;**2020**:5248569.

Brozek 2021

Brozek JL, Canelo-Aybar C, Akl EA, Bowen JM, Bucher J, Chiu WA, et al. GRADE guidelines 30: the GRADE approach to assessing the certainty of modeled evidence - An overview in the context of health decision-making. *Journal of Clinical Epidemiology* 2021;**129**:138-50.

Byambasuren 2020

Byambasuren O, Cardona M, Bell K, Clark J, McLaws M-L, Glasziou P. Estimating the extent of asymptomatic COVID-19 and its potential for community transmission: systematic review and meta-analysis. *medRxiv* 2020. [DOI: 10.1101/2020.05.10.20097543]

Cacciapaglia 2020a

Cacciapaglia G, Cot C, Sannino F. Second wave COVID-19 pandemics in Europe: a temporal playbook. *Scientific Reports* 2020;**10**(1):15514.

Cacciapaglia 2020b

Cacciapaglia G, Sannino F. Interplay of social distancing and border restrictions for pandemics via the epidemic renormalisation group framework. *Scientific Reports* 2020;**10**(1):15828.

Campbell 2020

Campbell M, McKenzie JE, Sowden A, Katikireddi SV, Brennan SE, Ellis S, et al. Synthesis without meta-analysis (SWiM) in systematic reviews: reporting guideline. *BMJ* 2020;**368**:l6890. [DOI: 10.1136/bmj.l6890]

Carfi 2020

Carfi A, Bernabei R, Landi F, for the Gemelli Against Covid-Post-Acute Care Study Group. Persistent symptoms in patients after acute COVID-19. *JAMA* 2020;**324**(6):603-5.

Caro 2014

Caro JJ, Eddy DM, Kan H, Kaltz C, Patel B, Eldessouki R, ISPOR-AMCP-NPC Modeling CER Task Forces. Questionnaire to assess relevance and credibility of modeling studies for informing health care decision making: an ISPOR-AMCP-NPC Good Practice Task Force report. *Value in Health* 2014;**17**(2):174-82. [DOI: 10.1016/j.jval.2014.01.003]

Chang 2020

Chang K, Pan CY, Lu PL. Sentinel surveillance at airports: Experience of dengue and COVID-19 prevention in Taiwan. *The Kaohsiung Journal of Medical Sciences* 2020;**36**(8):665-6. [DOI: 10.1002/kjm2.12265]

Chaudhry 2020

Chaudhry R, Dranitsaris G, Mubashir T, Bartoszko J, Riazi S. A country level analysis measuring the impact of government actions, country preparedness and socioeconomic factors on COVID-19 mortality and related health outcomes. *EClinicalMedicine* 2020;**25**:100464.

Chen 2020d

Chen M, Li M, Hao Y, Liu Z, Hu L, Wang L. The introduction of population migration to SEIAR for COVID-19 epidemic modeling with an efficient intervention strategy. *Information Fusion* 2020;**64**:252-8.

Dorjee 2020

Dorjee K, Kim H, Bonomo E, Dolma R. Prevalence and predictors of death and severe disease in patients hospitalized due to COVID-19: A comprehensive systematic review and meta-analysis of 77 studies and 38,000 patients. *PLoS One* 2020;**15**(12):e0243191. [DOI: 10.1371]

Egger 2017

Egger M, Johnson L, Althaus C, Schoni A, Salanti G, Low N, et al. Developing WHO guidelines: time to formally include evidence from mathematical modelling studies. *F1000Research* 2017;**6**:1584. [DOI: 10.12688/f1000research.12367.2]

Errett 2020

Errett NA, Sauer LM, Rutkow L. An integrative review of the limited evidence on international travel bans as an emerging infectious disease disaster control measure. *Journal of Emergency Management* 2020;**18**(1):7-14. [DOI: dx.doi.org/10.5055/jem.2020.0446]

Expert-Taskforce 2020

Expert-Taskforce. Epidemiology of COVID-19 Outbreak on Cruise Ship Quarantined at Yokohama, Japan, February 2020. Expert Taskforce for the Covid-Cruise Ship Outbreak. *Emerging infectious diseases* 2020;**26**(11):2591-7.

Fani 2020

Fani M, Teimoori A, Ghafari S. Comparison of the COVID-2019 (SARS-CoV-2) pathogenesis with SARS-CoV and MERS-CoV infections. *Future Virology* 2020;**10**(2217):fvl-2020-0050. [DOI: 10.2217/fvl-2020-0050]

Folayan 2015

Folayan M, Brown B. Ebola and the limited effectiveness of travel restrictions. *Disaster Medicine and Public Health Preparedness* 2015;**9**(1):92. [DOI: 10.1017/dmp.2015.1]

Furukawa 2020

Furukawa NW, Brooks JT, Sobel J. Evidence supporting transmission of severe acute respiratory syndrome coronavirus 2 while presymptomatic or asymptomatic. Emerging Infectious Diseases 2020;**26**(7). [DOI: 10.3201/eid2607.201595]

Garritty 2020

Garritty C, Gartlehner G, Kamel C, King VJ, Nussbaumer-Streit B, Stevens A, et al. Cochrane rapid reviews. Interim guidance from the Cochrane Rapid Reviews Methods Group (2020). methods.cochrane.org/rapidreviews/sites/methods.cochrane.org.rapidreviews/files/public/uploads/cochrane_rr_-_guidance-23mar2020-final.pdf.

Greenhalgh 2020

Greenhalgh T, Knight M, Court C, Buxton M, Husain L. Management of post-acute covid-19 in primary care. *BMJ* 2020;**370**:m3026.

Gupta 2020

Gupta A, Kunte R, Goyal N, Ray S, Singh K. A comparative analysis of control measures on-board ship against COVID-19 and similar novel viral respiratory disease outbreak: Quarantine ship or disembark suspects? Medical Journal Armed Forces India 2020 Jun 30 [Epub ahead of print].

Hayakawa 2020

Hayakawa K, Kutsuna S, Kawamata T, Sugiki Y Nonaka C, Tanaka K, et al. SARS-CoV-2 infection among returnees on charter flights to Japan from Hubei, China: a report from National Center for Global Health and Medicine. *Global Health & Medicine* 2020:2020.01036.

HHS 2020

US Department of Health & Human Services. What is the difference between isolation and quarantine? (2020). Available

at www.hhs.gov/answers/public-health-and-safety/what-is-the-difference-between-isolation-and-quarantine/index.html.

Higgins 2019

Higgins JP, Savović J, Page MJ, Elbers RG, Sterne JA. Chapter 8: Assessing risk of bias in a randomized trial. In: Higgins JP, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al, editor(s). Cochrane Handbook for Systematic Reviews of Interventions version 6.0 (updated July 2019). Cochrane, 2019. Available from www.training.cochrane.org/handbook.

Hilton Boon 2020

Hilton Boon M, Thomson H. The effect direction plot revisited: Application of the 2019 Cochrane Handbook guidance on alternative synthesis methods. *Research Synthesis Methods* 2020;**12**(1):29-33.

Hultcrantz 2017

Hultcrantz M, Rind D, Akl EA, Treweek S, Mustafa RA, Iorio A, et al. The GRADE Working Group clarifies the construct of certainty of evidence. *Journal of Clinical Epidemiology* 2017;**87**:4-13. [DOI: 10.1016/j.jclinepi.2017.05.006]

Ing 2020

Ing AJ, Cocks C, Green JP. COVID-19: in the footsteps of Ernest Shackleton. *Thorax* 2020;**75**(8):693-4.

Jablonska 2020

Jablonska K, Aballea S, Toumi M. Factors influencing the COVID-19 daily deaths peak across European countries. medRxiv 2020. [DOI: 10.1101/2020.11.04.20225656]

Jernigan 2020

Jernigan DB. Update: public health response to the coronavirus disease 2019 outbreak - United States, February 24, 2020. *Morbidity and Mortality Weekly Report, Surveillance Summaries : MMWR* 2020;**69**(8):216-9.

Jorritsma 2020

Jorritsma J, Hulshof T, Komjathy J. Not all interventions are equal for the height of the second peak. *Chaos, Solitons, and Fractals* 2020;**139**:109965.

Kiang 2020

Kiang MV, Chin ET, Huynh BQ, Chapman Lloyd AC, Rodriguez-Barraquer I, Greenhouse B, et al. Routine asymptomatic testing strategies for airline travel during the COVID-19 pandemic: a simulation analysis. *medRxiv* 2020:2020.12.08.20246132. [DOI: 10.1101/2020.12.08.20246132]

Koh 2020

Koh WC, Naing L, Wong J. Estimating the impact of physical distancing measures in containing COVID-19: an empirical analysis. *International Journal of Infectious Diseases* 2020;**100**:42-9.

Leffler 2020

Leffler CT, Ing E, Lykins JD, Hogan MC, McKeown CA, Grzybowski A. Association of country-wide coronavirus mortality with demographics, testing, lockdowns, and public

wearing of masks. *American Journal of Tropical Medicine and Hygiene* 2020;**103**(6):2153-4. [DOI: 10.4269/ajtmh.20-1342]

Liu 2020a

Liu Y, Morgenstern C, Kelly J, Lowe R, Jit M, CMMID COVID-19 Working Group. The impact of non-pharmaceutical interventions on SARS-CoV-2 transmission across 130 countries and territories. BMC Medicine 2021;**19**(1):40.

Moher 2009

Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA Statement. *PLoS Medicine* 2009;**6**(7):e1000097. [DOI: 10.1371/journal.pmed1000097]

Mouchtouri 2019

Mouchtouri VA, Christoforidou EP, An der Heiden M, Menel Lemos C, Fanos M, Rexroth U, et al. Exit and entry screening practices for infectious diseases among travelers at points of entry: looking for evidence on public health impact. *International Journal of Environmental Research & Public Health [Electronic Resource]* 2019;**16**(23):21. [DOI: dx.doi.org/10.3390/ijerph16234638]

Movsisyan 2021

Movsisyan A, Burns J, Biallas R. Travel-related control measures to contain the COVID-19 pandemic: an evidence map. BMJ Open (in press).

NICE 2020

National Institute for Health and Care Excellence. COVID-19 rapid guideline: managing the long-term effects of COVID-19. National Institute for Health and Care Excellence guideline [NG188]. 2020.

Nussbaumer-Streit 2020

Nussbaumer-Streit B, Mayr V, Dobrescu AI, Chapman A, Persad E, Klerings I, et al. Quarantine alone or in combination with other public health measures to control COVID-19: a rapid review. *Cochrane Database of Systematic Reviews* 2020, Issue 9.

Nuttal 2014

Nuttal I. Ebola travel: vigilance, not bans (commentary) World Health Organization. www.who.int/mediacentre/commentaries/ebola-travel/en/ (accessed prior to 22 March 2021).

Nuzzo 2014

Nuzzo JB, Cicero AJ, Waldhorn R, Inglesby TV. Travel bans will increase the damage wrought by ebola. *Biosecurity and Bioterrorism* 2014;**12**(6):306-9. [DOI: 10.1089/bsp.2014.1030]

Ogundokun 2020

Ogundokun RO, Lukman AF, Kibria GB, Awotunde JB, Aladeitan BB. Predictive modelling of COVID-19 confirmed cases in Nigeria. *Infectious Disease Modelling* 2020;**5**:543-8.

Ouzzani 2016

Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan-a web and mobile app for systematic reviews. *Systematic Reviews* 2016;**5**(1):210. [DOI: 10.1186/s13643-016-0384-4]

Cochrane
Library

Trusted evidence.
Informed decisions.
Better health.

Cochrane Database of Systematic Reviews

Philips 2006

Philips Z, Bojke L, Sculpher M, Claxton K, Golder S. Good practice guidelines for decision-analytic modelling in health technology assessment: a review and consolidation of quality assessment. *Pharmacoeconomics* 2006;**24**(4):355-71. [DOI: 10.2165/00019053-200624040-00006]

Potdar 2020

Potdar V, Choudhary ML, Bhardwaj S, Ghuge R, Sugunan AP, Gurav Y, et al. Respiratory virus detection among the overseas returnees during the early phase of COVID-19 pandemic in India. *Indian Journal of Medical Research* 2020;**151**(5):486-9.

Rees 2020

Rees EM, Nightingale ES, Jafari Y, Waterlow NR, Clifford S, Pearson CA, et al. COVID-19 length of hospital stay: a systematic review and data synthesis. BMC Medicine 2020;**18**(1).

Schünemann 2008

Schünemann HJ, Oxman AD, Brozek J, Glasziou P, Jaeschke R, Vist GE, GRADE Working Group. Grading quality of evidence and strength of recommendations for diagnostic tests and strategies. *BMJ* 2008;**336**(7653):1106-10. [DOI: 10.1136/bmj.39500.677199.AE]

Schünemann 2019

Schünemann HJ, Cuello C, Akl EA, Mustafa RA, Meerpohl JJ, Thayer K, GRADE Working Group. GRADE guidelines: 18. How ROBINS-I and other tools to assess risk of bias in nonrandomized studies should be used to rate the certainty of a body of evidence. *Journal of Clinical Epidemiology* 2019;**111**:105-14. [DOI: 10.1016/j.jclinepi.2018.01.012]

Sriwijitalai 2020a

Sriwijitalai W, Wiwanitkit V. Positive screening for Wuhan novel coronavirus infection at international airport: what's the final diagnosis for positive cases. *International Journal of Preventive Medicine* 2020a;**11**:30.

Sterne 2016

Sterne JA, Hernan MA, Reeves BC, Savovic J, Berkman ND, Viswanathan M, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ* 2016;**355**:i4919. [DOI: 10.1136/bmj.i4919]

Stokes 2020

Stokes J, Turner AJ, Anselmi L, Morciano M, Hone T. The relative effects of non-pharmaceutical interventions on early Covid-19 mortality: natural experiment in 130 countries. *medRxiv* 2020. [DOI: 10.1101/2020.10.05.20206888]

Swadi 2021

Swadi T, Geoghegan J, Devine T, McElnay C, Sherwood J, Shoemack P, et al. Genomic evidence of in-flight transmission of SARS-CoV-2 despite predeparture testing. *Emerging Infectious Diseases* 2021;**27**(3):687-93.

Teixeira da Silva 2020

Teixeira da Silva JA, Tsigaris P. Policy determinants of COVID-19 pandemic-induced fatality rates across nations. *Public Health* 2020;**187**:140-2.

Tognotti 2013

Tognotti E. Lessons from the history of quarantine, from plague to influenza A. *Emerging Infectious Diseases* 2013;**19**(2):254-9. [DOI: 10.3201/eid1902.120312]

Viswanathan 2020

Viswanathan M, Kahwati L, Jahn B, Giger K, Dobrescu AI, Hill C, et al. Universal screening for SARS-CoV-2 infection: a rapid review. *Cochrane Database of Systematic Reviews* 2020, Issue 9.

Whiting 2011

Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, QUADAS-2 Group. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Annals of Internal Medicine* 2011;**155**(8):529-36. [DOI: 10.7326/0003-4819-155-8-201110180-00009]

WHO 2005

World Health Organization. International Health Regulations. 2nd edition. Geneva: World Health Organization, 2005.

WHO 2019

World Health Organization. Non-pharmaceutical public health measures for mitigating the risk and impact of epidemic and pandemic influenza. Geneva: World Health Organization, 2019.

WHO 2020a

World Health Organization. Novel Coronavirus (2019-nCoV) Situation Report - 1. Available from www.who.int/docs/default-source/coronaviruse/situation-reports/20200121-sitrep-1-2019-ncov.pdf?sfvrsn=20a99c10_4 2020.

WHO 2020b

World Health Organization. Novel Coronavirus (2019-nCoV) Situation Report - 51. Available from apps.who.int/iris/bitstream/handle/10665/331475/nCoVsitrep11Mar2020-eng.pdf?sequence=1&isAllowed=y 2020.

Wu 2020

Wu Z, McGoogan JM. Characteristics of and important lessons from the coronavirus disease 2019 (COVID-19) outbreak in China: summary of a report of 72314 cases from the Chinese Center for Disease Control and Prevention. *JAMA* 2020;**323**(13):1239-42. [DOI: 10.1001/jama.2020.2648]

References to other published versions of this review

Burns 2020

Burns J, Movsisyan A, Stratil JM, Coenen M, Emmert-Fees KM, Geffert K, et al. Travel-related control measures to contain the COVID-19 pandemic: a rapid review. *Cochrane Database of Systematic Reviews* 2020, Issue 10. Art. No: CD013717. [DOI: 10.1002/14651858.CD013717]

* Indicates the major publication for the study

CHARACTERISTICS OF STUDIES

Characteristics of included studies *[ordered by study ID]*

Adekunle 2020

Study characteristics	
Study design	Mathematical modelling study • Description: stochastic meta-population SEIR model * Global stochastic meta-population SEIR model with two infectious stages (both symptomatic, one with lower infectiousness) using migration patterns based on international flight travel volumes to estimate the impact of a travel ban.
Disease	COVID-19
Travel-related control measure(s)	Border closure • Wuhan lockdown and travel ban on China • Travel ban on Iran, South Korea and Italy Date of implementation: different travel bans implemented progressively since January 24, 2020
Country implementing the measure(s)	Country protected by the measure: Australia Country restricted by the measure: China, Iran, South Korea, and Italy
Outcome(s)	Cases avoided due to the measure • Outcome: number of imported cases avoided * Follow-up: 1 December - 24 March, 2020
Notes	COI: "The authors have stated they have no conflict of interest." Funding: n.r.

Al-Qahtani 2020

Study characteristics	
Study design	Observational screening study
Disease	COVID-19
Travel-related control measure(s)	Screening at national borders and quarantine of travellers • Entry screening procedures: PCR-based testing upon arrival and final PCR-test at the end of the 14-day period prior to discharge • Quarantine procedures: 14-day mandatory quarantine with PCR-based testing upon developing symptoms of a respiratory infection Date of implementation: 25 February 2020
Country implementing the measure(s)	Country protected by the measure: Kingdom of Bahrain Country restricted by the measure: n.r.

Al-Qahtani 2020 *(Continued)*

Outcome(s)	Cases detected due to the measure • Outcome: number of cases identified in index test vs. reference test * Follow-up: 18 days; quarantine and observation period was 14 days
Notes	COI: "The authors have declared that no conflict of interest exists." Funding: "No funding was received to perform this study."

Al-Tawfiq 2020

Study characteristics	
Study design	Observational screening study
Disease	COVID-19
Travel-related control measure(s)	Screening at national borders and quarantine of travellers • Entry screening procedures: symptom-based screening/triage at day 0; PCR-based testing for all arrivals; re-testing with rt-PCR at day 12/13 prior to discharge • Quarantine procedures: quarantine for 14 days with regular symptom-based screening and PCR-testing upon development of symptoms Date of implementation: 14 March 2020
Country implementing the measure(s)	Country protected by the measure: Saudi Arabia Country restricted by the measure: USA, United Kingdom, Spain, Italy, Canada, Egypt, Dubai, Oman and Bahrain
Outcome(s)	Cases detected due to the measure • Outcome: number of cases identified in index test vs. reference test * Follow-up: the observation period is 10 weeks (first arrival in facility on 14 Match, last arrival on 4 May); quarantine and observation period was 14 days
Notes	COI: "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." Funding: "There is no funding for this study."

Anderson 2020

Study characteristics	
Study design	Mathematical modelling study • Description: extended Bayesian SEIR model * Complex Bayesian SEIR model with a compartment for quarantine and two exposed stages (non-infectious and infectious). Includes modelling of a fixed proportion of the population participating in social distancing behaviour and is estimated separately for each jurisdiction.
Disease	COVID-19

Anderson 2020 *(Continued)*

Travel-related control measure(s)	Border closure • Relaxation of the closure of national borders • Date of implementation: not specified
Country implementing the measure(s)	Country or region protected by the measure: California; Sweden; Ontario; Washington; United Kingdom; Quebec; British Columbia; New York; Germany; Belgium; New Zealand; Japan Country restricted by the measure: n.r.
Outcome(s)	Cases avoided due to the measure • Outcome: number of additional cases introduced when one infected traveller enters • Follow-up: 6 weeks after relaxation
Notes	COI: n.r. Funding: "The authors have stated they have no conflict of interest"

Anzai 2020

Study characteristics	
Study design	Mathematical modelling study • Description: counterfactual projection model based on Poisson process * The model assumes a Poisson process determining exported cases to destination countries from China and the probability of a major epidemic in destination countries based on a negative binomial distribution of generated secondary cases. This counterfactual projection is compared to observed exported cases. Model allows for untraced cases and imperfect contact tracing.
Disease	COVID-19
Travel-related control measure(s)	Multiple measures • Complete border closure • Travel restrictions • Quarantine of travellers • Entry screening for all incoming travellers Date of implementation: January 23 2020
Country implementing the measure(s)	Country protected by the measure: n.r./Japan for outcome 2 Country restricted by the measure: China

Outcome(s)	Cases avoided due to the measure
	
	• Outcome 1: number of exported out cases from China
	  * Follow-up: 28 January – 6 February 2020
	
	Shift in epidemic development
	
	• Outcome 2: absolute risk reduction in the probability of major epidemic
	  * Follow-up: 28 January – 6 February 2020
	
	Shift in epidemic development

Anzai 2020 *(Continued)*

- Outcome 3: time to a major epidemic
  * Follow-up: 28 January 28 – 6 February 2020

Notes	COI: "The authors declare no conflicts of interest."
	Funding: "H.N. received funding from the Japan Agency for Medical Research and Development (AMED) [grant number: JP18fk0108050]; the Japan Society for the Promotion of Science (JSPS) KAKENHI [grant numbers, H.N.: 17H04701, 17H05808, 18H04895 and 19H01074; R.K.: 18J21587; AS.: 19K24159], the Inamori Foundation, and the Japan Science and Technology Agency (JST) CREST program [grant number: JPMJCR1413]. SMJ and NML receive graduate study scholarships from the Ministry of Education, Culture, Sports, Science and Technology, Japan."

Arima 2020

Study characteristics	
Study design	Observational screening study
Disease	COVID-19
Travel-related control measure(s)	Screening at national borders and quarantine of travellers (evacuation flight)
	• Entry screening procedures: likely-symptom focused clinical examination; rt-PCR of oro-pharyngeal swab samples
	* Reference test for entry and/or exit screening: observation of cases during quarantine; PCR test for all those who developed symptoms during quarantine; rt-PCR test of oropharyngeal swab samples on day 1; and at the end of the 14-day quarantine period for all asymptomatic passengers
	• Quarantine procedures: quarantine of all passengers tested negative for 14 days
	* Reference test for quarantine observation: PCR test for all those who developed symptoms during quarantine; rt-PCR test of oropharyngeal swab samples on day 1; and at the end of the 14-day quarantine period for all asymptomatic passengers
	Date of implementation: 29-31 January 2020
Country implementing the measure(s)	Country protected by the measure: Tokyo, Japan
	Country restricted by the measure: Hubei Province, China
Outcome(s)	Cases detected due to the measure
	• Outcome 1: proportion of cases detected (symptom screening)
	• Outcome 2: positive predictive value (symptom screening)
	• Outcome 3: proportion of cases detected (PCR screening)
	• Outcome 4: positive predictive value (PCR screening)
	• Outcome 5: proportion of cases detected (combined symptom and PCR screening)
	• Outcome 6: positive predictive value (combined symptom and PCR screening)
	* Follow-Up: 29-31 January 2020 with follow-up reference testing
Notes	COI: n.r.
	Funding: "This study was supported in part by a grant-in-aid from the Japan Agency for Medical Research and Development (grant nos. JP19fk0108104 and JP19fk0108110)."

Ashcroft 2020

Study characteristics

Study design	Mathematical modelling study • Description: unspecified (mathematical model using empirical distributions of transmission parameters) * Relative utility of quarantine strategies for contacts and returning travellers under different assumptions, contact tracing and test result delays is estimated. Relative utility is defined as the ratio between the amount of overall transmission prevented and the number of person days spent in quarantine.
Disease	COVID-19

Travel-related control measure(s)	Quarantine of travellers
	
	• Quarantine of travellers returning from high risk countries
	  * Days of quarantine: 1, 2, 3, 5, 7, 10, and 14
	• Quarantine of travellers with test-release
	  * Test on arrival and release on day 2
	  * Test on day 3 and release on day 5
	  * Test on day 5 and release on day 7

Country implementing the measure(s)	Country protected by the measure: not specified Country restricted by the measure: not specified
Outcome(s)	Cases detected due to the measure • Outcome: proportion of cases detected * Follow-up: time spent in quarantine (up to 10 days)
Notes	COI: n.r. Funding: n.r.

Banholzer 2020

Study characteristics

Study design	Mathematical modelling study • Description: Bayesian hierarchical model * The number of new cases (modeled based on a negative binomial distribution) is linked to the number of existing cases, country and time parameters and the presence of any NPIs (assumed to have the same effectiveness in each country). The model then, under a counterfactual scenario, estimates the relative reduction in new cases.
Disease	COVID-19
Travel-related control measure(s)	Border closure • Closure of national borders for individuals Date of implementation: varied in 12 countries implementing the measure
Country implementing the measure(s)	Country protected by the measure: Austria, Australia, Belgium, Canada, Denmark, France, Finland, Italy, Norway, Switzerland, Spain, United States

Banholzer 2020 *(Continued)*

	Country restricted by the measure: all other countries
Outcome(s)	Cases avoided due to the measure • Outcome: relative reduction in number of new cases * Follow-up: through 15 April 2020
Notes	COI: "SF reports further grants from the Swiss National Science Foundation outside of the submitted work. JPS declares part-time employment at Luciole Medical outside of the submitted work. All other authors declare no competing interests." Funding: "NB, EvW and SF acknowledge funding from the Swiss National Science Foundation (SNSF) as part of the Eccellenza grant 186932 on 'Data-driven health management.'"

Bays 2020

Study characteristics

Study design	Mathematical modelling study • Description: mechanistic mathematical model simulating individual travel using Monte Carlo methods
Disease	COVID-19
Travel-related control measure(s)	Screening at national borders and quarantine of travellers • Entry screening of all arriving travellers • Entry testing with quarantine of all travellers * Days of quarantine: 3, 5, 7, 10, and 14 Date of implementation: not reported
Country implementing the measure(s)	Country protected by the measure: n.r. Country restricted by the measure: n.r.
Outcome(s)	Cases detected due to the measure • Outcome: proportion of cases detected * Follow-up: n.r.
Notes	COI: n.r. Funding: n.r.

Binny 2020

Study characteristics

Study design	Mathematical modelling study

Binny 2020 *(Continued)*

- Description: continuous-time stochastic branching process model of COVID-19 transmission and control developed for New Zealand
  * Initial seed cases represent overseas arrivals replicating real case data.
  * At each time step (day) individuals produce a Poisson distributed number of secondary infections with the mean corresponding to an equation of transmission parameters.
  * Transmission parameters are based on theoretical and empirical distributions.
  * Interventions are modeled based on change in transmission parameters.
  * Several alternative timing scenarios and components of New Zealand's strategy are compared.

Disease	COVID-19
Travel-related control measure(s)	Border closure and quarantine of travellers • Applied to all except returning citizens and residence • Mandatory home quarantine of all international arrivals for 14 days Date of implementation: 11 - 15 March 2020
Country implementing the measure(s)	Country protected by the measures: New Zealand Countries restricted by the measures: all other countries

Outcome(s)	Cases avoided due to the measure
	
	 • Outcome 1: maximum number of daily new reported cases
	 • Outcome 2: cumulative number of cases
	   * Follow-up: 10 March - 12 May 2020
	
	 Shift in epidemic development
	
	 • Outcome 3: number of daily cases at peak
	 • Outcome 4: probability of eliminating epidemic
	   * Follow-up: 10 March - 12 May 2020

Notes	COI: n.r. Funding: n.r.

Boldog 2020

Study characteristics	

Study design	Mathematical modelling study
	
	 • Description: time-dependant SEIR model as input for global transportation network model and Galton-Watson branching process in destination.
	   * A time-dependent SEIR model is used to model transmission dynamics and estimate the cumulative number of cases, which is then used as an input parameter for a global transportation network model which generates probability distributions of the number of exported cases at each destination. Finally, a Galton-Watson branching process in each destination country estimates the probability of a major outbreak.
	   * Gamma distributed incubation and infectious period based on SARS-study.

Disease	COVID-19
Travel-related control measure(s)	International travel restrictions and entry screening

Boldog 2020 *(Continued)*

	Date of implementation: n.r.
Country implementing the measure(s)	Country protected by the measure: USA, Canada, Thailand and South Korea
	Country restricted by the measure: China
Outcome(s)	Shift in epidemic development
	• Outcome: risk of major outbreak
	* Follow-up: n.r.
Notes	COI: "The authors declare no conflict of interest".
	Funding: "G.R. was supported by EFOP-3.6.1-16-2016-00008. F.B. was supported by NKFIH KKP 129877. T.T. was supported by NKFIH FK 124016. A.D. was supported by NKFIH PD 128363 and by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences. P.B. was supported by 20391-3/2018/FEKUSTRAT".

Chen J 2020

Study characteristics

Study design	Observational screening study
Disease	COVID-19
Travel-related control measure(s)	Screening at national borders
	• Screening procedures: all passengers were required to take their temperature before deplaning; deplaning was followed by 15 or 22 day quarantine with daily symptom-assessment and PCR-based testing upon development of symptoms; PCR-based test of all individuals +2 days after arrival; PCR-based test of all individuals + 13 days after arrival
	Date of implementation: 24 January 2020
Country implementing the measure(s)	Country protected by the measure: China, Hangzhou
	Country restricted by the measure: China, Singapore
Outcome(s)	Cases detected due to the measure
	• Outcome: number of cases identified in index test vs. reference test
	* Follow-up: 15 or 21 days, depending on the risk profile
Notes	COI: "For all authors none were declared."
	Funding: "This work was financially supported by grants from Zhejiang province (Zhejiang Scientific and Technological Major Project) under the 2020 Emergency (Grant No. 2020C03124), Zhejiang University special scientific research fund for COVID-19 prevention and control (Grant No. 2020XGZX047) and Technology Project of Hangzhou Municipality (Grant No. 20202013A02)."

Chen T 2020

Study characteristics

Study design	Mathematical modelling study

Chen T 2020 *(Continued)*

	• Description: modified deterministic SEIR model
Disease	COVID-19
Travel-related control measure(s)	Quarantine of travellers and entry restrictions • Quarantine of travellers crossing national borders • Entry restrictions at national borders for different intensity of restriction * Target group: all inbound passengers and symptomatic travellers only Date of implementation: 13 March 2020: new visitors from Italy, France, Germany and Spain were not allowed entry into Singapore; 23 March 2020: all short-term visitors were prohibited from entering or transition through Singapore
Country implementing the measure(s)	Country protected by the measure: China, Singapore Countries restricted by the measure: all other countries
Outcome(s)	Cases avoided due to the measure • Outcome 1: number of cases in community • Outcome 2: number of imported cases * Follow-up: 21 May 2020
Notes	COI: "We declare no competing interests" Funding: "None."

Chen Y-H 2020

Study characteristics	
Study design	Mathematical modelling study • Description: modified deterministic SEIR model
Disease	COVID-19
Travel-related control measure(s)	Quarantine of travellers with test-and isolation, contact tracing, and general public mask-wearing/so-cial-distancing Date of implementation: 31 December, 2019
Country implementing the measure(s)	Country protected by the measure: Taiwan Countries restricted by the measure: n.r.
Outcome(s)	Cases avoided due to the measure • Outcome 1: number of cases and hospitalized cases in 90 days • Outcome 2: number quarantined in the community in 90 days * Follow-up: n.r.
Notes	COI: "The authors have no conflicts of interest relevant to this article." Funding: "The authors acknowledge the financial support provided by Ministry of Health and Welfare and National Taiwan University Infectious Diseases Research and Education Center, Taipei, Taiwan.

Chen Y-H 2020 *(Continued)*

The funders have no role in study design, data collection and analysis, or preparation of the manuscript."

Chinazzi 2020

Study characteristics

Study design	Mathematical modelling study
	• Description: individual-based, stochastic, and spatial meta-population epidemic model
	* The global epidemic and mobility model (GLEAM) uses a meta-population network approach which divides the real-world population in subpopulations centered around transportation hubs. Ground and air travel mobility flows are estimated from real-world data and transmission dynamics are estimated in each subpopulation using a SEIR model. Model assumes the detection of imported cases to not be lower than 40% and travel probabilities, susceptibility and contact patterns to be homogenous.
Disease	COVID-19
Travel-related control measure(s)	International travel restrictions
	• International travel restrictions on China, including suspension and limitation of flights to and from China
	Date of implementation: Wuhan travel ban implemented on 23 January 2020; China travel restrictions implemented on 1 February 2020.
Country implementing the measure(s)	Country protected by the measure: n.r.
	Country restricted by the measure: China
Outcome(s)	Cases avoided due to the measure
	• Outcome: number of imported cases
	* Follow-up: 1 January – 1 March 2020
Notes	COI: "M.E.H. reports grants from the National Institute of General Medical Sciences during the conduct of the study; A.V. reports grants and personal fees from Metabiota, Inc., outside of the submitted work; M.C. and A.P.yP. report grants from Metabiota, Inc., outside of the submitted work; H.Y. reports grants from Glaxosmithkline (China) Investment Co., Ltd., Yichang HEC Changjiang Pharmaceutical Co., Ltd, Sanofi Pasteur, and Shanghai Roche Pharmaceuticals Company, outside of the submitted work. The authors declare no other relationships or activities that could appear to have influenced the submitted work."
	Funding: "M.E.H. acknowledges the support of the MIDASU54GM111274. S.M. and M.A. acknowledge support from the EU H2020 MOOD project. C.G. and L.R. acknowledge support from the EU H2020 Icarus project. M.C. and A.V. acknowledge support from Google Cloud Healthcare and Life Sciences Solutions via the GCP research credits program. The findings and conclusions in this study are those of the authors and do not necessarily represent the official position of the funding agencies, the National Institutes of Health, or the U.S. Department of Health and Human Services."

Clifford 2020a

Study characteristics

Study design	Mathematical modelling study

Clifford 2020a *(Continued)*

- Description: poisson process and local outbreak model
  * The model integrates a stochastic non-homogeneous Poisson process estimating arrival of infected travellers with a local outbreak probability model where interventions increase probability of identification of infected cases or reduce the reproductive number and thus reduce outbreak probability. Individuals with severe symptoms are assumed not to attempt travel.

Disease	COVID-19
Travel-related control measure(s)	Screening at borders Date of implementation: n.r.
Country implementing the measure(s)	Country protected by the measure: n.r. Country restricted by the measure: n.r.
Outcome(s)	Shift in epidemic development - Outcome: delay of outbreak - Follow-up: n.r.
Notes	COI: "We declare no competing interests" Funding: "SF and SC are supported by a Sir Henry Dale Fellowship jointly funded by the Wellcome Trust and the Royal Society (Grant number 208812/Z/17/Z). RME acknowledges an HDR UK Innovation Fellowship (Grant number MR/S003975/1). BQ was funded by the National Institute for Health Research (NIHR) (16/137/109) using UK aid from the UK Government to support global health research. PK was funded by the Royal Society under award RP\EA\180004 and by the Bill & Melinda Gates Foundation (INV-003174). KvZ is supported by Elrha's Research for Health in Humanitarian Crises (R2HC) Programme, which aims to improve health outcomes by strengthening the evidence base for public health interventions in humanitarian crises. The R2HC programme is funded by the UK Government (DFID), the Wellcome Trust, and the UK National Institute for Health Research (NIHR). CABP gratefully acknowledges funding by the Department for International Development / Wellcome Epidemic Preparedness Coronavirus research programme (ref. 221303/Z/20/Z) and by the NTD Modelling Consortium by the Bill and Melinda Gates Foundation (OPP1184344. The following funding sources are acknowledged as providing funding for the working group authors. Alan Turing Institute (AE). BBSRC LIDP (BB/M009513/1: DS). This research was partly funded by the Bill & Melinda Gates Foundation (INV-003174: KP, MJ, YL; NTD Modelling Consortium OPP1184344: GM; OPP1180644: SRP; OPP1183986: ESN; OPP1191821: KO'R, MA). ERC Starting Grant (#757688: CJVA, KEA; #757699: JCE, RMGJH). This project has received funding from the European Union's Horizon 2020 research and innovation programme - project EpiPose (101003688: KP, MJ, WJE, YL). This research was partly funded by the Global Challenges Research Fund (GCRF) project 'RECAP' managed through RCUK and ESRC (ES/P010873/1: AG, CIJ). Nakajima Foundation (AE). This research was partly funded by the National Institute for Health Research (NIHR) using UK aid from the UK Government to support global health research (16/137/109: CD, FYS, MJ, YL; Health Protection Research Unit for Modelling Methodology HPRU-2012-10096: NGD, TJ; PR-OD-1017-20002: AR). The views expressed in this publication are those of the author(s) and not necessarily those of the NIHR or the UK Department of Health and Social Care. RCUK/ESRC (ES/P010873/1: TJ). Royal Society (Dorothy Hodgkin Fellowship: RL). UK DHSC/UK Aid/NIHR (ITCRZ 03010: HPG). UK MRC (LID DTP MR/N013638/1: EMR, QJL; MR/P014658/1: GMK). Authors of this research receive funding from UK Public Health Rapid Support Team funded by the United Kingdom Department of Health and Social Care (TJ). Wellcome Trust (206250/Z/17/Z: AJK, TWR; 210758/Z/18/Z: JDM, JH, NIB, SA, SFunk, SRM). No funding (AKD, AMF, DCT, SH"

Clifford 2020b

Study characteristics

Clifford 2020b *(Continued)*

Study design	Mathematical modelling study
	• Description: mathematical model estimating empirical distributions of transmission parameters
	* Potential traveller trajectories under different quarantine and screening scenarios are calculated
	* Scenarios are compared based on number of infectious travellers released and number of infectious days per released traveller
Disease	COVID-19
Travel-related control measure(s)	Screening at national borders and quarantine of travellers
	• Procedures: quarantine for various lengths; multiple rounds of PCR testing: pre-flight (mandatory in some countries), post-flight, and follow-up
	Date of implementation: n.r.
Country implementing the measure(s)	Country protected by the measure: United Kingdom
	Country restricted by the measure: EU, USA
Outcome(s)	Cases detected due to the measure
	• Outcome 1: number of infectious travellers being released in the community
	• Outcome 2: number of remaining infectious days for each infected traveller after their release from quarantine
	* Follow-up: ip to 14 days
Notes	COI: "Akira Endo received a research grant from Taisho Pharmaceutical Co. Ltd."
	Funding: "The following funding sources are acknowledged as providing funding for the named authors. This research was partly funded by the Bill & Melinda Gates Foundation (INV-003174: YL; NTD Modelling Consortium OPP1184344: CABP). DFID/Wellcome Trust (Epidemic Preparedness Coronavirus research programme 221303/Z/20/Z: CABP). This project has received funding from the European Union's Horizon 2020 research and innovation programme - project EpiPose (101003688: WJE, YL). HDR UK (MR/S003975/1: RME). This research was partly funded by the National Institute for Health Research (NIHR) using UK aid from the UK Government to support global health research. The views expressed in this publication are those of the author(s) and not necessarily those of the NIHR or the UK Department of Health and Social Care (16/136/46: BJQ; 16/137/109: BJQ, YL; PR-OD-1017-20002: WJE). UK MRC (MC_PC_19065 - Covid 19: Understanding the dynamics and drivers of the COVID-19 epidemic using real-time outbreak analytics: RME, SC, WJE, YL). Wellcome Trust (206250/Z/17/Z: TWR; 208812/Z/17/Z: SC, SFlasche). Alan Turing Institute and Nakajima Foundation (AE). No funding (YWDC)."

Costantino 2020

Study characteristics	
Study design	Mathematical modelling study
	• Description: poisson process and age-specific, deterministic extended SEIR model
	* The model assumes a Poisson process to estimate the possible true epidemic curve in China and then calculates the number of infected entering the country. An age-specific deterministic compartmental (susceptible (S), latent traced (Et), latent untraced (Eu), infectious (I), isolated (Q), recovered (R) and dead (D)) is then used to estimate transmission dynamics due to imported cases. The proportion of asymptomatic infections is assumed to be 34.6% and notified cases reflect only 10% of real infections per day.
Disease	COVID-19

Costantino 2020 *(Continued)*

Travel-related control measure(s)	International travel restrictions/bans • No travel ban; complete travel ban followed by complete lifting of the ban; complete travel ban followed by partial lifting of the ban (allowing university students, but not tourists, to enter the country) Date of implementation: 1 February 2020
Country implementing the measure(s)	Country protected by the measure: Australia Country restricted by the measure: China
Outcome(s)	Cases avoided due to the measure • Outcome 1: number of exported cases from China • Outcome 2: number of cases in Australia • Outcome 3: number of deaths in Australia * Follow-up: 400 days Shift in epidemic development • Outcome 4: delay of the epidemic outbreak • Follow-up: 400 days
Notes	COI: "The authors declare no conflicts of interest." Funding: "This work has no funding."

Davis 2020

Study characteristics	
Study design	Mathematical modelling study • Description: established individual-based, stochastic, and spatial epidemic model: Global Epidemic and Mobility Model (GLEAM) * Global population is divided into 3200 sub-populations in 200 countries/territories. * Transmission dynamics in subpopulation assume a SLIR model. * Individuals transition between compartments through stochastic chain binomial processes.
Disease	COVID-19
Travel-related control measure(s)	International travel restrictions Date of implementation: 2 February 2020
Country implementing the measure(s)	Country protected by the measure: USA Country restricted by the measure: China
Outcome(s)	Shift in epidemic development • Outcome: time to onset of local transmission (defined as 100 new infections per day) * Follow-up: 5 January - 5 March 2020
Notes	COI: "M.E.H. reports grants from National Institute of General Medical Sciences, during the conduct of the study; A.V. reports grants and personal fees from Metabiota inc., outside the submitted work; M.C. and A.P.P. report grants from Metabiota inc., outside the submitted work. No other relationships or activities that could appear to have influenced the submitted work."

Davis 2020 *(Continued)*

Funding: "M.E.H. reports grants from National Institute of General Medical Sciences, during the conduct of the study; A.V. reports grants and personal fees from Metabiota inc., outside the submitted work; M.C. and A.P.P. report grants from Metabiota inc., outside the submitted work. No other relationships or activities that could appear to have influenced the submitted work. M.E.H. reports grants from National Institute of General Medical Sciences, during the conduct of the study; A.V. reports grants and personal fees from Metabiota inc., outside the submitted work; M.C. and A.P.P. report grants from Metabiota inc., outside the submitted work. No other relationships or activities that could appear to have influenced the submitted work."

Deeb 2020

Study characteristics	
Study design	Mathematical modelling study
	• Description: SEIR model of situation in Lebanon adapted to account for incoming travellers
	* Time-varying R(t) represents local social distancing measures in place in Lebanon
	* Various rates of increased incoming travellers used to predict effect of relaxation of travel restriction
Disease	COVID-19
Travel-related control measure(s)	International travel restriction
	• Airport closure
	Date of implementation: n.r.
Country implementing the measure(s)	Country protected by the measure: Lebanon
	Country restricted by the measure: n.r.
Outcome(s)	Cases avoided due to the measure
	• Outcome: cumulative number of cases
	* Follow-up: 130 days
Notes	COI: "The authors declare that they have no conflict of interests."
	Funding: n.r.

Dickens 2020

Study characteristics	
Study design	Mathematical modelling study
	• Description: mathematical model using empirical distributions of transmission parameters to simulate travellers from various countries
	* Measures are assessed through the number of simulated passengers meeting certain conditions
Disease	COVID-19
Travel-related control measure(s)	Test-based screening at national borders, quarantine of travellers and entry restrictions
	• Entry screening procedures: entry testing of all incoming travellers

Dickens 2020 *(Continued)*

	• Quarantine procedures: quarantine of different lengths • Entry restrictions: prohibiting entry for those testing positive Date of implementation: 23 July 2020
Country implementing the measure(s)	Country protected by the measure: n.r. Country restricted by the measure: n.r.
Outcome(s)	Cases avoided due to the measure • Outcome 1: number of imported cases • Outcome 2: number of cases in community * Follow-up: n.r.
Notes	COI: "None declared." Funding: "The work was supported by funding from Singapore's National Medical Research Council through grants COVID19RF-004 and NMRC/CG/C026/2017_NUHS."

Gostic 2020

Study characteristics	
Study design	Mathematical modelling study • Description: unspecified probabilistic model * Same as Gostic 2015 with COVID-19-specific disease parameters and extensions pertaining to ways in which infected travellers can be missed, fraction of subclinical cases, and the incubation period distribution.
Disease	COVID-19
Travel-related control measure(s)	Screening at national borders • Hypothetical thermal infrared fever scanner and exposure-focused questionnaire at departure, arrival, and departure and arrival Date of implementation: n.r.
Country implementing the measure(s)	Country protected by the measure: n.r. Country restricted by the measure: n.r.
Outcome(s)	Cases detected due to the measure • Outcome: proportion of infected travellers detected * Follow-up: n.r.
Notes	COI: n.r. Funding: the study mentions the following funders: James S. McDonnell Foundation for funding K.G. (postdoctoral fellowship in dynamic and multi-scale systems), Wellcome for funding A.J.K. (206250/Z/17/Z), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior for A.C.R. G. (science without-borders fellowship), National Science Foundation for J.O.L-S (DEB-1557022), Defense Advanced Research projects Agency for J.O.L-S (PREEMPT D18AC00031), and Strategic Environmental Research and-Development Program for A.C.R.G, R.O.M, and J.O.L-S (RC-2635). "The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication."

Grannell 2020

Study characteristics

Study design	Mathematical modelling study
	
	 • Description: two region SEIR model for the Island of Ireland
	   * Standard four compartments are used (susceptible, exposed, infected, and recovered)
	   * Model accounts for interaction between the border regions of Ireland and Northern Ireland

Disease	COVID-19
Travel-related control measure(s)	Border closure • Closing of border from the first day of the epidemic for international travelers Date of implementation: first day of the epidemic
Country implementing the measure(s)	Country protected by the measure: Ireland and Northern Ireland Country restricted by the measure: Ireland and Northern Ireland
Outcome(s)	Shift in epidemic development due to the measure • Outcome 1: days to epidemic peak • Outcome 2: Proportion of population infected at peak * Follow-up: n.r.
Notes	COI: n.r. Funding: n.r.

Hoehl 2020

Study characteristics

Study design	Observational screening study
Disease	COVID-19

Travel-related control measure(s)	Screening at national borders and quarantine of travellers (evacuation flight)
	
	 • Entry screening procedures: likely-symptom focused clinical examination
	 • Exit screening procedures: symptoms of COVID-19; contact with an infected person, accompanying a person with symptoms or with a contact
	   * Reference test for entry/exit screening: observation of cases during quarantine (no further details provided)l PCR test for all passengers (once, likely upon arrival)
	 • Quarantine procedures: quarantine of all passengers tested negative for 14 days
	   * Reference test for quarantine observation: PCR test for all passengers (once, likely upon arrival)
	
	 Date of implementation: 1 February 2020

Country implementing the measure(s)	Country protected by the measure: Germany Country restricted by the measure: Hubei, China
Outcome(s)	Cases detected due to the measure

Hoehl 2020 *(Continued)*

- Outcome 1: proportion of cases detected
- Outcome 2: positive predictive value
  - Follow up: 1 day (one time event) with follow-up reference testing

Notes	COI: n.r.
	Funding: n.r.

James 2020

Study characteristics	
Study design	Mathematical modelling study
	- Description: model-free estimation of COVID-19 transmission dynamics
Disease	COVID-19
Travel-related control measure(s)	Border closure and quarantine of travellers
	Date of implementation: 16 March 2020: 14-day quarantine; 20 March 2020: border closure to non-citizens and non-residents, 10 April 2020: 14-day mandatory government-managed quarantine
Country implementing the measure(s)	Country protected by the measure: New Zealand
	Country restricted by the measure: n.r.
Outcome(s)	Cases avoided due to the measure
	- Outcome 1: number of secondary infections caused by an imported case - Outcome 2: probability of an imported case not causing any further infection - Follow-up: n.r.
Notes	COI: n.r.
	Funding: "This work was funded by the Ministry of Business, Innovation and Employment and Te Pūnaha Matatini, New Zealand's Centre of Research Excellence in complex systems."

Kang 2020

Study characteristics	
Study design	Modelling study

	- Description: synthetic control design
	  * Uses outcome data and key confounders from the intervention countries with travel bans and from non-intervention countries to construct a 'synthetic' country as counterfactual
	  * 39 countries were part of the non-intervention donor pool

Disease	COVID-19
Travel-related control measure(s)	Border closure
	Date of implementation: Australia: 1 February 2020; Singapore: 2 February 2020; US: 2 February 2020; Vietnam: 3 February 2020; Taiwan: 7 February 2020; Hong Kong: 8 February 2020

Kang 2020 *(Continued)*

Country implementing the measure(s)	Country protected by the measure: Australia; Singapore; US; Vietnam; Taiwan; Hong Kong
	Country restricted by the measure: China
Outcome(s)	Cases avoided due to the measure
	• Outcome: cumulative number of cases
	* Follow-up: until 29 February 2020
Notes	COI: "The authors have no conflicts of interest associated with the material presented in this paper."
	Funding: "None."

Kim 2020

Study characteristics	
Study design	Observational screening study
Disease	COVID-19
Travel-related control measure(s)	Screening at national borders and quarantine of travellers
	• Screening procedures: fever-based symptom screening and symptom-based health questionnaire pre-departure; infection-protection measures on the airplane (e.g. physical distancing, hygiene concept); fever-based symptom screening and symptom-based health questionnaire upon arrival; rt-PCR test for all arrivals
	• Quarantine of travellers for 14 days
	Date of implementation: n.r.
Country implementing the measure(s)	Country protected by the measure: South Korea
	Country restricted by the measure: Iran
Outcome(s)	Cases detected due to the measure
	• Outcome: number of cases identified in index test vs. reference test
	* Follow-up: one day event with a 14-day quarantine observation
Notes	COI: "The authors have no conflicts of interest to declare."
	Funding: n.r.

Kivuti-Bito 2020

Study characteristics	
Study design	Mathematical modelling study
	• Description: stock-flow model
	* Modified, deterministic SEIR Model
Disease	COVID-19

Kivuti-Bito 2020 *(Continued)*

Travel-related control measure(s)	Quarantine of travellers (compulsory)
	Date of implementation: 13 March 2020 (one month after first case in Kenya)
Country implementing the measure(s)	Country protected by the measure: Kenya
	Countries restricted by the measure: all other countries
Outcome(s)	Shift in epidemic development
	• Outcome 1: days to peak in in active cases
	• Outcome 2: days to peak in daily cases and deaths
	* Follow-up: n.r.
Notes	COI: "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."
	Funding: "This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors."

Kwok 2020

Study characteristics	
Study design	Mathematical modelling study
	• Description: SEIR Model with different localized patches with travel in-between
	* Covers Hong-Kong dynamics with travel from China
	* No specific changes in dynamical equations
	* Temperature dependent R0 controls dynamics
Disease	COVID-19
Travel-related control measure(s)	Border closure
	• Reduction in number of daily travellers from 200,000 to 0
	Date of implementation: 8 February 2020
Country implementing the measure(s)	Country protected by the measure: Hong Kong, China
	Country restricted by the measure: China, excluding Hubei
Outcome(s)	Cases avoided due to the measure
	• Outcome 1: Cumulative number of cases
	• Outcome 2: number of deaths
	* Follow-up: 8 February - 14, June 2020
Notes	COI: "No potential conflict of interest was reported by the authors."
	Funding: "This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors."

Lagier 2020

Study characteristics

Study design	Observational screening study
Disease	COVID-19
Travel-related control measure(s)	Screening at national borders • Screening procedures: symptom-based screening upon arrival (and during the flight); rt-PCR test around 24h after arrival (day 3 to 8 of quarantine); rt-PCR retest of all individuals on day 5 after arrival and day 7 to 13 after starting the quarantine • Quarantine procedures: quarantine of arrivals from China for 2-7 days (procedures not described); quarantine with observation for the development of symptoms approximately 14 days after arrival. Date of implementation: n.r.
Country implementing the measure(s)	Country protected by the measure: France Country restricted by the measure: Wuhan, China
Outcome(s)	We reported no outcomes, as no cases were identified in the study.
Notes	COI: n.r. Funding: "This work has benefited from the French State support, managed by the 'Agence Nationale de la Recherche' including the "Programme d'Investissement d'avenir" under the reference.Méditer-ranée Infection 10-IAHU-03 Funding are no any involvement in the study design; collection, analysis and interpretation of data; the writing of the manuscript."

Liebig 2020

Study characteristics

Study design	Mathematical modelling study
	
	• Description: unspecified
	   * Expected number of arrivals into Australia assuming no travel restrictions were implemented are predicted by fitting a seasonal autoregressive integrated moving average model on data between January 2015 and December 2019
	   * Ascertainment rates are estimated by fitting a Binomial distribution to the number of infected in-dividuals among the arrivals into Australia
	   * The expected number of importations into Australia is calculated based on the length of overseas stay of a traveller and the daily incidence rates of COVID-19 in the country

Disease	COVID-19
Travel-related control measure(s)	Border closure and its relaxation Date of implementation: 1 February - 20 March 2020: different levels of travel ban; October 2020: relax-ation.
Country implementing the measure(s)	Country protected by the measure: Australia Country restricted by the measure: UK, Italy, New Zealand, France, Austria, Switzerland, US, China, Netherlands
Outcome(s)	Cases avoided due to the measure

Liebig 2020 *(Continued)*

	• Outcome: proportion of imported cases
	* Follow-up: 1 January - 30 June 2020
Notes	COI: "The authors declare no competing interests."
	Funding: "This work is part of the DiNeMo project." European Commission

Linka 2020a

	Study characteristics
Study design	Mathematical modelling study
	• Description: SEIR model combined with mobility network model
	* Per country discretized SEIR model based on a network graph representation of Europe to integrate travel and transmission dynamics and estimate daily increments in each compartment in each country using mobility coefficients.
Disease	COVID-19
Travel-related control measure(s)	Border closure
	Date of implementation: 17 March 2020
Country implementing the measure(s)	Country protected by the measure: 27 countries of the EU
	Country restricted by the measure: n.r.
Outcome(s)	Cases avoided due to the measure
	• Outcome 1: proportion of infectious individuals in the population
	* Follow-up: 23 March - 20 April 2020
	Shift in epidemic development
	• Outcome 2: time from introduction of travel restriction until inflection point in R(t)
	* Follow-up: R(t) tracked from introduction of specific travel measure until inflection point
Notes	COI: "The authors declare no conflict of interest."
	Funding: "This work was supported by a Stanford Bio-X IIP seed grant to Mathias Peirlinck and Ellen Kuhl and by a DAAD Fellowship to Kevin Linka."

Linka 2020b

	Study characteristics
Study design	Mathematical modelling study
	• Description: SEIR model combined with mobility network model
	* Discretised SEIR model based on a network graph representation of North America to integrate travel and transmission dynamics and estimate daily increments in each compartment in Newfoundland and Labrador
Disease	COVID-19

Trusted evidence.
Informed decisions.
Better health.

Cochrane Database of Systematic Reviews

Linka 2020b *(Continued)*

Travel-related control measure(s)	Border closure Date of implementation: 1 July 2020
Country implementing the measure(s)	Country protected by the measure: Canada, Newfoundland/Labrador Country restricted by the measure: USA
Outcome(s)	Shift in epidemic development • Outcome: days to infection of 0.1 percent population * Follow-up: 150 days
Notes	COI: "The authors declare that they have no conflict of interest." Funding: "This work was supported by a DAAD Fellowship to Kevin Linka, the Engineering and Physical Sciences Research Council grant EP/R020205/1 to Alain Goriely, and a Stanford Bio-X IIP seed grant and the National Institutes of Health Grant U01 HL119578 to Ellen Kuhl."

Lio 2020

Study characteristics

Study design	Observational screening study
Disease	COVID-19
Travel-related control measure(s)	Screening at national borders and quarantine of travellers • Entry screening procedures: rt-PCR of nasopharyngeal swabs on day 2; rt-PCR of nasopharyngeal swabs on day 7 and day 13 for viral RNA; assessment of Sera antibodies on day 14 before citizens are released from quarantine • Quarantine procedures: 14-day quarantine with symptom-observation following arrival; Date of implementation: 7 March 2020
Country implementing the measure(s)	Country protected by the measure: Macao, China Country restricted by the measure: Wuhan, China
Outcome(s)	We reported no outcomes, as no cases were identified in the study.
Notes	COI: "The authors declare that they have no competing interests." Funding: "The authors received no funding for this work."

Lytras 2020

Study characteristics

Study design	Observational screening study
Disease	COVID-19

International travel-related control measures to contain the COVID-19 pandemic: a rapid review (Review)

68

Copyright © 2021 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Lytras 2020 *(Continued)*

Travel-related control measure(s)	Screening at national borders and quarantine of travellers (evaluation flights)
	
	• Entry screening procedures: symptom-based arrival screening
	  * Reference test for entry screening: PCR-testing of passengers upon arrival; observation of passengers for 14 days (without additional PCRs)
	• Quarantine procedures: Quarantine of all passengers tested negative for 14 days
	  * Reference test for quarantine observation: none
	
	Date of implementation: 22-25 March 2020

Country implementing the measure(s)	Country protected by the measure: Greece Country restricted by the measure: UK, Spain, Turkey
Outcome(s)	Cases detected due to the measure • Outcome 1: proportion of cases detected • Outcome 2: positive predictive value * Follow-up: n.r.
Notes	COI: "None declared." Funding: "No funding was received for this study."

Mandal 2020

Study characteristics	
Study design	Mathematical modelling study • Description: extended SEIR model * A deterministic epidemic is simulated with an SEIR model to inform projections for daily cases introduced at points of entry in the destination country based on a Binomial distribution. Symptomatic and asymptomatic cases are assumed to have different travel probabilities and some proportion of both is detected during entry screening.
Disease	COVID-19
Travel-related control measure(s)	Screening at national borders • Entry screening of symptomatic passengers arriving from China at ports of entry Date of implementation: n.r.
Country implementing the measure(s)	Country protected by the measure: India Country restricted by the measure: China
Outcome(s)	Shift in epidemic development • Outcome: number of days to reach a prevalence of 1000 cases * Follow-up: n.r.
Notes	COI: "None" Funding: "None"

McLure 2020

Study characteristics

Study design	Mathematical modelling study
	
	• Description: modified version of the model proposed by Costantino 2020. Poisson process and age-specific, deterministic extended SEIR model
	  * An age-specific deterministic compartmental (susceptible (S), latent traced (Et), latent untraced (Eu), infectious (I), isolated (Q), recovered (R) and dead (D)) is used to estimate transmission dynamics due to imported cases.
	  * The proportion of asymptomatic infections is assumed to be 34.6% and notified cases reflect only 10% of real infections per day.
	  * Modification: Excluded cases from Hubei after the severe lockdown of Hubei starting on 23 January as this lockdown made it very unlikely that citizens from Hubei travelled to Australia

Disease	COVID-19
Travel-related control measure(s)	International travel restrictions/bans • No travel ban; complete travel ban followed by complete lifting of the ban; complete travel ban followed by partial lifting of the ban (allowing university students, but not tourists, to enter the country) Date of implementation: 1 February 2020
Country implementing the measure(s)	Country protected by the measure: Australia Country restricted by the measure: China
Outcome(s)	Cases avoided due to the measure • Outcome: number of imported cases * Follow-up: 26 January - 4 April 2020
Notes	COI: "None declared." Funding: "Australian National Health and Medical Research Council Early Career Fellowships (APP1158469 to L.F.K.)."

Nakamura 2020

Study characteristics

Study design	Mathematical modelling study • Description: unspecified * Calculation of risk of importation and exportation of COVID-19 between 1491 airports based on a passenger flow matrix proposed by Huang et al. (2010, Plos One), data on confirmed COVID-19 cases until March 14 2020, provided by WorldPop and a risk flow calculation proposed by Gilbert et al. (2020, The Lancet)
Disease	COVID-19
Travel-related control measure(s)	International travel restrictions • Reductions of air travel by 90% (for the airports in the 1st quartile area with regard to cumulative COVID-19 incidence in airport area), 60% (for the airports in the 2nd quartile) and 30% (for the airports in the 3rd quartile)

Nakamura 2020 *(Continued)*

	Date of implementation: n.r.
Country implementing the measure(s)	Country protected by the measure: n.r.
	Country restricted by the measure: n.r.
Outcome(s)	Cases avoided due to the measure
	• Outcome: risk of importation and exportation * Follow-up: n.r.
Notes	COI: "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."
	Funding: "This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors."

Ng 2020

Study characteristics	
Study design	Observational screening study
Disease	COVID-19
Travel-related control measure(s)	Screening at national borders and quarantine of travellers (evacuation flights)

	• Screening procedures: fever-based arrival and departure screening
	  * Reference test for screening: medical assessment of quarantined 3x/day; PCR assessment of those showing symptoms; PCR-essay on day 3 and day 6 of quarantine of most of those quarantined
	• Quarantine procedures: Quarantine of all passengers tested negative for 14 days
	  * Reference test for quarantine: medical assessment of quarantined 3x/day; PCR assessment of those showing symptoms; PCR-essay on day 3 and day 6 of quarantine of most of those quarantined

	Date of implementation: 30 January 2020
Country implementing the measure(s)	Country protected by the measure: Singapore
	Country restricted by the measure: Wuhan, China
Outcome(s)	Cases detected due to the measure
	• Outcome 1: proportion of cases detected • Outcome 2: positive predictive value * Follow-up: 1 day (one time event) with follow-up reference testing
Notes	COI: n.r.
	Funding: "Supported by the following grants from the National Medical Research Council, Ministry of Health, Singapore: Collaborative Solutions Targeting Antimicrobial Resistance Threats in Health Systems (CoSTAR-HS) (NMRC CGAug16C005), NMRC Clinician Scientist Award (MOH-000276), and NMRC Clinician Scientist Individual Research Grant (MOH-CIRG18nov-0006)."

Nowrasteh 2020

Study characteristics

Study design	Modelling study
	• Description: synthetic control design
	* Uses outcome data and key confounders from the USA and from countries without a ban on travel from China to construct a 'synthetic' country as counterfactual
	* 13 countries were part of the non-intervention donor pool
Disease	COVID-19
Travel-related control measure(s)	Border closure
	• Banning the entry of all aliens who were physically present in China during the 14-day period preceding their entry or attempted entry into the United States, with some exceptions for U.S. lawful permanent residents and those closely related to American citizens
	Date of implementation: 2 February 2020
Country implementing the measure(s)	Country protected by the measure: USA
	Country restricted by the measure: China
Outcome(s)	Cases avoided due to the measure
	• Outcome 1: cumulative number of cases
	• Outcome 2: cumulative cases per million
	• Outcome 3: number of new cases
	• Outcome 4: number of new cases per million
	* Follow-up: 22 January - 9 March 2020
Notes	COI: n.r.
	Funding: n.r.

Nuckchady 2020

Study characteristics

Study design	Mathematical modelling study
	• Description: modified stochastic SEIR model
Disease	COVID-19
Travel-related control measure(s)	International travel restrictions
	• Reduction of infected passengers to 1 per 100 days
	Date of implementation: n.r.
Country implementing the measure(s)	Country protected by the measure: Mauritius
	Country restricted by the measure: n.r.
Outcome(s)	Shift in epidemic development
	• Outcome 1: days to epidemic outbreak

Nuckchady 2020 *(Continued)*

	• Outcome 2: probability of outbreak * Follow-up: n.r.
Notes	COI: n.r.
	Funding: n.r.

Odendaal 2020

Study characteristics

Study design	Mathematical modelling study • Description: unspecified * Data-driven approach. The model uses observed data to fit an exponential curve and project the cumulative number of infected into the future. Asymptomatic infections are assumed to become contagious some time after the moment of infection from exposure and the average person who has been infected will show symptoms after the average incubation period.
Disease	COVID-19
Travel-related control measure(s)	Border closure Date of implementation: 31 January 2020
Country implementing the measure(s)	Country protected by the measure: USA Country restricted by the measure: China
Outcome(s)	Shift in epidemic development • Outcome: delay of community transmission • Follow-up: 3 months
Notes	COI: n.r. Funding: n.r.

Pinotti 2020

Study characteristics

Study design	Mathematical modelling study
	
	• Description: mathematical model using empirical distributions of transmission parameters
	  * Gamma distribution is fitted to model detection delay as a function of the day the case arrived at destination
	  * Case arrival accounts for detected and undetected cases. The expected number of imported cases follows a Poisson distribution with an exponential growth parameter which varies with the location of origin and with containment measures in place in the location of origin
	  * Index case detection model: multinomial model in which the number of observed clusters with imported index cases, the number of known imported cases causing no onward transmission is compared to the number of cluster for which no index case was identified to estimate the number of undetected imported cases causing no onward transmission (latent variable) and the case detection probability

Pinotti 2020 *(Continued)*

Disease	COVID-19
Travel-related control measure(s)	International travel restrictions • Travel restrictions were modeled by assuming a discontinuity in the growth rate Date of implementation: 23 January 2020: Hubei; 29 January 2020: rest of China
Country implementing the measure(s)	Country protected by the measure: n.r. Country restricted by the measure: Hubei Province; China
Outcome(s)	Cases avoided due to the measure • Outcome: exportation growth rate * Follow-up: 5 January - 27 February 2020
Notes	COI: "The authors have declared that no competing interests exist." Funding: "This study is partially funded by the Agence National de la Recherche (ANR, https://anr.fr/) through the project DATAREDUX (ANR-19-CE46-0008-03) to VC; the European Union with grants RECOV-ER (H2020-101003589) and MOOD (H2020-874850, https://ec.europa.eu/programmes/horizon2020/en) to PYB, CP, and VC; REACTing (https://reacting.inserm.fr/) through the COVID-19 funding to VC; the Municipality of Paris (https://www.paris.fr/) through the programme Emergence(s) to FP and CP; INSERM-INRIA partnership for research on public health and data science to LDD. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript."

Quilty 2020

Study characteristics	
Study design	Mathematical modelling study • Description: unspecified (Microsimulation) * Microsimulation of travellers planning to board a flight with simulated disease histories and different travel durations, screening detection probabilities and proportions of asymptomatic cases. Condition of infection worsens after some time, barring travellers from flying.
Disease	COVID-19
Travel-related control measure(s)	Screening at national borders • Fever screening with thermal image scanners (sensitivity 86%) on departure and arrival Date of implementation: n.r.
Country implementing the measure(s)	Country protected by the measure: n.r. Country restricted by the measure: n.r.
Outcome(s)	Cases detected due to the measure • Outcome: proportion of infected travellers detected * Follow-up: n.r.
Notes	COI: "None." Funding: "SF and SC are supported by a Sir Henry Dale Fellowship jointly funded by the Wellcome Trust and the Royal Society (Grant number 208812/Z/17/Z). RME acknowledges an HDR UK Innova-

Quilty 2020 *(Continued)*

tion Fellowship (Grant number MR/S003975/1). BJQ was funded by the National Institute for Health Research (NIHR) (16/137/109) using UK aid from the UK Government to support global health research. The views expressed in this publication are those of the author(s) and not necessarily those of the NHS, the NIHR or the UK Department of Health and Social Care. CMMID nCoV working group funding statements: Yang Liu (Gates (INV-003174), NIHR (16/137/109)), Charlie Diamond (NIHR (16/137/109)), Sebastian Funk (Wellcome Trust (210758/Z/18/Z)), Amy Gimma (Global Challenges Research Fund (GCRF) for the project "RECAP" managed through RCUK and ESRC (ES/P010873/1)), James D Munday (Wellcome Trust (210758/Z/18/Z)), Hamish Gibbs (NIHR (ITCRZ 03010)), Sam Abbott (Wellcome Trust (210758/Z/18/Z)), Timothy W Russell (Wellcome Trust (206250/Z/17/Z)), Petra Klepac (Gates (INV-003174)), Mark Jit (Gates (INV-003174), NIHR (16/137/109)), Joel Hellewell (Wellcome Trust (210758/Z/18/Z))."

Russell TW 2020

Study characteristics

Study design	Mathematical modelling study
	
	• Description: comparison of total incidence and imported cases
	   * Imported cases via international air travel considered
	   * Prevalences were adjusted for country-specific under-ascertainment

Disease	COVID-19
Travel-related control measure(s)	International travel restrictions, border closure or quarantine of travellers • Defined as any measure that completely or almost completely prevents international arrivals from contributing to local transmission, such as entry bans and compulsory 14-day facility-based quarantines Date of implementation: n.r.
Country implementing the measure(s)	Country protected by the measure: multiple countries (142) Country restricted by the measure: n.r.
Outcome(s)	Cases avoided due to the measure • Outcome: proportion of imported cases in overall cases * Follow-up: the month of May 2020
Notes	COI: "We declare no competing interests." Funding: "Wellcome Trust, UK Department for International Development, European Commission, National Institute for Health Research, Medical Research Council, Bill & Melinda Gates Foundation."

Russell WA 2020

Study characteristics

Study design	Mathematical modelling study
	
	• Description: unspecified
	   * Mathematical model using empirical distributions of transmission parameters to simulate travellers from various countries
	   * Measures are assessed through the number of simulated passengers meeting certain conditions

Russell WA 2020 *(Continued)*

Disease	COVID-19
Travel-related control measure(s)	Screening at national borders and quarantine of travellers • Entry screening procedures: entry testing (no quarantine) • Quarantine procedures: mandatory quarantine of arriving travellers without testing and mandatory quarantine of arriving travellers with testing Date of implementation: n.r.
Country implementing the measure(s)	Country protected by the measure: n.r. Country restricted by the measure: n.r.
Outcome(s)	Cases detected due to the measure • Outcome: days at risk of transmission * Follow-up: 14 days
Notes	COI: "The authors have no conflicts to disclose." Funding: "WAR was supported by a Stanford Interdisciplinary Graduate Fellowship; DLB was supported by a Canada Research Chair in Health Informatics and Data Science."

Ryu 2020

Study characteristics	
Study design	Mathematical modelling study • Description: extended deterministic SEIR model * SEIR model including quarantine compartment and additional pre-infection seeding as influx into the exposed class to estimate daily infection counts over time under quarantine scenarios. Assumes that there is no ongoing community transmission at the destination and that there is no pre-symptomatic infectiousness.
Disease	COVID-19
Travel-related control measure(s)	Quarantine of travellers • All quarantined individuals were confined at home or to the university dormitory as per the current South Korean quarantine program for COVID-19 (the model assessed scenarios with different rates of quarantine compliance: 70%, 80%, 90% or 100%) Date of implementation: 15 days before and after 1 March 2020
Country implementing the measure(s)	Country protected by the measure: South Korea Country restricted by the measure: China
Outcome(s)	Cases avoided due to the measure • Outcome: total number of infected individuals seeded by arriving students * Follow-up: 180 days
Notes	COI: "The authors declare no conflict of interest." Funding: "This research received no external funding."

Cochrane
Library

Trusted evidence.
Informed decisions.
Better health.

Cochrane Database of Systematic Reviews

Shaikh Abdul Karim 2020

Study characteristics

Study design	Observational screening study
Disease	COVID-19
Travel-related control measure(s)	Screening at national borders and quarantine of travellers
	• Screening procedures: multiple testing-related measures prior and after the repatriation flight: pre-flight medical screening/assessment; post-arrival symptom-based health screening with PCR testing for all individuals;
	• Quarantine procedures: quarantine and surveillance 14 days post arrival with PCR on day 13 post discharge
	Date of implementation: five repatriation missions were performed between February and the end of April 2020
Country implementing the measure(s)	Country protected by the measure: Malaysia
	Country restricted by the measure: China, Iran, Italy, Indonesia
Outcome(s)	Cases detected due to the measure
	• Outcome: number of cases identified in index test vs. reference test
	* Follow-up: second PCR test was performed on day 13 after arrival
Notes	COI: "The author(s) declare(s) that they have no competing interests."
	Funding: n.r.

Shi 2020

Study characteristics

Study design	Mathematical modelling study
	• Description: survival Time analysis of time until county detects first case
	* Country-specific, time-constant hazard function
	* Spread by air travel from Wuhan, characterized by concept of effective distance
Disease	COVID-19
Travel-related control measure(s)	Border closure and international travel restrictions
	• Defined as entry or exit travel bans for travellers to/from China; visa restrictions as total or partial suspensions for travellers from China; flight suspensions as governmental bans to or from China. Travel restrictions reduced 25%, 50%, and 75% of flights from China to countries in which restrictions were in place.
	Date of implementation: varied across 80 countries
Country implementing the measure(s)	Country protected by the measure: multiple countries (80)
	Country restricted by the measure: China
Outcome(s)	Cases avoided due to the measure

Shi 2020 *(Continued)*

	• Outcome: risk of case importation * Follow-up: 8 December, 2019 - 26 February 2020
Notes	COI: "None declared." Funding: n.r.

Sruthi 2020

Study characteristics	
Study design	Mathematical modelling study • Description: data -driven AI-approach * Counts of reported cases in Swiss cantons used to estimate weekly infection rates * Effect of different contributions of non-pharmaceutical interventions on reproduction number
Disease	COVID-19
Travel-related control measure(s)	Border closure • Full restriction of travel in and out; partial relaxation - land-border to Schengen countries opened Date of implementation: full restrictions: March 2020; partial relaxation: middle of June 2020
Country implementing the measure(s)	Country protected by the measure: Switzerland Country restricted by the measure: all other countries; non-Schengen countries.
Outcome(s)	Shift in epidemic development • Outcome: time varying reproduction rate * Follow-up: 9 March - 13 September 2020
Notes	COI: "The authors declare no competing interests." Funding: n.r.

Steyn 2020

Study characteristics	
Study design	Mathematical modelling study • Description: continuous-time branching process to mode. * Mathematical model to describe the status of disease (e.g. time since infection, symptomatic-status) within a simulated population
Disease	COVID-19
Travel-related control measure(s)	Screening at national borders and quarantine of travellers • Screening procedures: exit and entry testing of all travellers without quarantine • Quarantine: quarantine of all arriving travellers with or without testing: 5-day quarantine with one test on day 3; 14-day quarantine with no scheduled testing; 14-day quarantine with two tests on days 3 and 12

Steyn 2020 *(Continued)*

	Date of implementation: n.r.
Country implementing the measure(s)	Country protected by the measure: New Zealand
	Country restricted by the measure: n.r.
Outcome(s)	Cases avoided due to the measure
	• Outcome 1: probability of infected case released into community * Follow-up: n.r.
Notes	COI: n.r.
	Funding: "This work was funded by the New Zealand Ministry of Business, Innovation and Employment and Te Pūnaha Matatini, Centre of Research Excellence in Complex Systems."

Taylor 2020

Study characteristics	
Study design	Mathematical modelling study

	• Description: reduction of potentially infectious people by probabilistically stacking different measures
	  * Custom method developed in paper
	  * Applied to UK airports

Disease	COVID-19
Travel-related control measure(s)	Screening at national borders and quarantine of travellers
	• Entry screening procedures: thermal imaging scanner; health checks; single PCR testing (4 days after arrival); double PCR testing (upon arrival and after 4 days); single PCR testing (7 days after arrival); double PCR testing (upon arrival and after 7 days) • Quarantine procedures: quarantine for 7, 10, and 14 days
	Date of implementation: n.r.
Country implementing the measure(s)	Country protected by the measure: United Kingdom
	Country restricted by the measure: top-25 countries flying commercially to the UK (making up 86% of international flights into UK airports)
Outcome(s)	Cases avoided due to the measure
	• Outcome 1: proportion of infectious arrivals prevented from onward transmission * Follow-up: one week in August 2020
	Cases detected due to the measure
	• Outcome 2: proportion of detected infectious passengers * Follow-up: n.r.
Notes	COI: "None declared."
	Funding: "Public Health England (PHE) commissioned and funded this work as well as providing data on worldwide cases, deaths and tests for COVID-19. They did not have any role in the model design, analysis or preparation of the manuscript."

Utsunomiya 2020

Study characteristics

Study design	Mathematical modelling study
	• Description: moving Regression (MR) and Hidden Markov Model (HMM)
	* Data-driven approach. Framework for the real-time decomposition of infection growth curves into growth rate and acceleration and classification of stages ("lagging", "exponential", "deceleration", "stationary") which can be used to track intervention effects over time. Assumes that changes in cases per day are solely attributable to intervention.
Disease	COVID-19
Travel-related control measure(s)	International travel restrictions
	• International travel restrictions as defined by the Oxford COVID-19 Government Response Tracker (Ox-CGRT)
	Date of implementation: n.r.
Country implementing the measure(s)	Country protected by the measure: multiple countries (62)
	Country restricted by the measure: n.r.
Outcome(s)	Shift in epidemic development
	• Outcome: average change in epidemic growth acceleration after intervention implementation
	* Follow-up: n.r.
Notes	COI: "The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest."
	Funding: "This study did not receive financial support and was conducted during voluntary social isolation."

Wells 2020

Study characteristics

Study design	Mathematical modelling study
	• Description: custom model based on Maximum Likelihood Methods
	* Daily probability that an infected person travels abroad is fitted to observed data and the distribution of incubation period is used to predict importation to other countries weighted by international flight data. Time to the first transmission event was then estimated using the distribution of the serial interval and the incubation period.
Disease	COVID-19
Travel-related control measure(s)	Border closure
	• Lockdown of Wuhan and 15 other cities in Hubei
	Date of implementation: 23 January 2020: Wuhan; 24 January 2020: other cities in China
Country implementing the measure(s)	Country protected by the measure: all other countries

Wells 2020 *(Continued)*

	Country restricted by the measure: China

Outcome(s)	Cases avoided due to the measure
	
	• Outcome 1: number of exported cases from China
	    * Follow-up: 6 December, 2019 - 15 February 2020
	
	Cases detected due to the measure
	
	• Outcome 2: number of cases exported from China detected by airport screening
	    * Follow-up: 6 December, 2019 - 15 February 2020

Notes	COI: the authors declare no competing interest Funding: n.r.

Wilson 2020

Study characteristics	
Study design	Mathematical modelling study • Description: stochastic SEIR model * Published stochastic SEIR model is used but its application in the context of simulated flights and details on modeling of in-inflight transmission as well as outbreak modeling in destination country is not disclosed. Passengers and crew members arrive daily at disease-free destinations from countries with low prevalence. Passengers and crew can be in various stages of infection and cause infections on flight and in New Zealand.
Disease	COVID-19
Travel-related control measure(s)	Screening at national borders and quarantine of travellers • Procedures: various combinations of symptom based exit screening; symptom based entry screening; PCR tests on (3.1) day 1, (3.2) day 1 + 8, (3.3) day 3-12; contact tracing between PCR test; self-reporting of symptoms; quarantine for (6.1) 7 days or (6.2) 14 days Date of implementation: n.r.
Country implementing the measure(s)	Country protected by the measure: New Zealand (hypothetical disease free area) Country restricted by the measure: Australia (hypothetical low prevalence area)
Outcome(s)	Shift in epidemic development • Outcome: time to outbreak * Follow-up: n.r.
Notes	COI: "The authors have no competing interests." Funding: "Professor Wilson is supported by the New Zealand Health Research Council (16/443) and Ministry of Business Innovation and Employment (MBIE) funding of the BODE3 Programme (UOOX1406). Professor Michael Baker is supported by a New Zealand Health Research Council grant for research on COVID-19 (20/1066)."

Wong J 2020

Study characteristics

Study design	Observational screening study
Disease	COVID-19
Travel-related control measure(s)	Screening at national borders • Entry screening procedures: rt-PCR based testing of all arrivals • Quarantine procedures: 14-day quarantine of all arrivals Date of implementation: 21 March 2020
Country implementing the measure(s)	Country protected by the measure: Brunei Country restricted by the measure: all other countries
Outcome(s)	Cases detected due to the measure • Outcome: number of cases identified in index test vs. reference test * Follow-up: n.r.
Notes	COI: "None declared." Funding: n.r.

Wong MC 2020

Study characteristics

Study design	Mathematical modelling study
Disease	COVID-19
Travel-related control measure(s)	Quarantine of travellers • Compulsory 14-day quarantine at designated camps or at home of travellers crossing national borders Date of implementation: 30 January - 4 February 2020: 1st wave; 25 February - 25 March 2020: 2nd wave
Country implementing the measure(s)	Country protected by the measure: China Country restricted by the measure: Italy, Japan, Europe, all countries
Outcome(s)	Shift in epidemic development • Outcome: cumulative cases in the community * Follow-up: 8 January - 25 February 2020: 1st wave; 26 February - 24 April 2020: 2nd wave
Notes	COI: "None declared." Funding: "The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors."

Yamahata 2020

Study characteristics

Study design	Observational screening study
Disease	COVID-19
Travel-related control measure(s)	Screening at national borders and quarantine of travellers (Diamond Princess cruise ship) • Entry screening procedures: screening based on symptoms of COVID-19 and being a close contact of someone with symptoms • Quarantine procedures: Quarantine on board until first PCR test was conducted (2 to 23 days). Quarantine for 14 days of all those who tested negative * Reference test: testing of all individuals on board with PCR over a 3 week period after arrival; quarantine and observation on board until first PCR test; quarantine and observation for 14 days after a negative PCR test; PCR testing of those under quarantine showing symptoms Date of Implementation: 5 February 2020 - 23 February 2020
Country implementing the measure(s)	Country protected by the measure: Japan Country restricted by the measure: n.r.
Outcome(s)	Cases detected due to the measure • Outcome 1: proportion of cases detected • Outcome 2: positive predictive value * Follow-up: 1 day (one time event) with follow-up reference testing
Notes	COI: "None declared." Funding: "The authors did not receive any funding for this study"

Yang 2020

Study characteristics

Study design	Mathematical modelling study • Description: SEIR Model describing the case data internationally as one global system • Includes mobility data, social distancing, case isolation * Stochastic approach with time dependent parameters
Disease	COVID-19
Travel-related control measure(s)	Border closure • Shutdown in Wuhan; complete international travel ban were executed from different time points Date of implementation: January - October 2020 with various simulations using different start dates
Country implementing the measure(s)	Country protected by the measure: multiple countries: Hubei, China, US, Switzerland, Sweden, Austria, France, UK, Germany, Spain, Italy, Netherlands, Belgium, Denmark Country: restricted by the measure: all other countries
Outcome(s)	Cases avoided due to the measure

Yang 2020 *(Continued)*

	• Outcome: number of daily cases
	* Follow-up: January - October 2020
Notes	COI: "The authors declare no competing interests."
	Funding: "Funding was provided by National Natural Science Foundation of China (Grant Nos. 71974011, 71972012,71804009) and Beijing Social Science Fund (Grant Nos. 17JDGLB008, 17GLC043). Special Fund for Joint Development Program of Beijing Municipal Commission of Education."

Zhang C 2020

Study characteristics

Study design	Mathematical modelling study
	• Description: linear model incorporating effect of travel restrictions
	* Travel between 22 countries
	* Also takes into account internal movement restrictions
Disease	COVID-19
Travel-related control measure(s)	Border closure and international travel restrictions
	• International travel restriction and entry ban to reduce human mobility between countries (varying in the 22 countries)
	Date of implementation: 22 January - 24 April 2020
Country implementing the measure(s)	Country protected by measure: Multiple countries (22: USA, Spain, Italy, France, Germany, UK, Turkey, Iran, China, Russia, Brazil, Belgium, Canada, Netherlands, Switzerland, India, Portugal, Ecuador, Japan, South Korea, Australia, South Africa)
	Country restricted by measure: all other countries
Outcome(s)	Cases avoided due to the measure
	• Outcome: daily new cases
	* Follow-up: 22 January - 24 April 2020
Notes	COI: n.r.
	Funding: n.r.

Zhang L 2020

Study characteristics

Study design	Mathematical modelling study
	• Description: index for case importation risk on international flights is established
Disease	COVID-19
Travel-related control measure(s)	International travel restrictions
	• All inbound international flights were cancelled

**Cochrane
Library**

**Trusted evidence.
Informed decisions.
Better health.**

Cochrane Database of Systematic Reviews

Zhang L 2020 *(Continued)*

	Date of implementation: 29 March 2020
Country implementing the measure(s)	Country protected by the measure: China
	Country restricted by the measure: all other countries
Outcome(s)	Cases avoided due to the measure
	• Outcome: imported case risk index * Follow-up: 4 March - 16 April 2020
Notes	COI: "We also do not see any conflict of interest involved in this submission."
	Funding: "We also acknowledge gratefully the financial support from National Social Science Foundation of China (18ZDA071) and National Natural Science Foundation of China (71901065)."

Zhong 2020

Study characteristics	
Study design	Mathematical modelling study
Disease	COVID-19
Travel-related control measure(s)	Border closure and international travel restrictions
	• Radical travel restrictions of different levels (i.e. entry ban, global travel ban, and lockdown)
	Date of implementation: 21 January - 04 April 2020
Country implementing the measure(s)	Country protected by the measure: multiple countries (249 geographic areas)
	Country restricted by the measure: n.r.
Outcome(s)	Cases avoided due to the measure
	• Outcome 1: number of cumulative cases in community • Outcome 2: days to epidemic arrival * Follow-up: 21 January - 4 April 2020
Notes	COI: "The authors declare that they have no competing financial interests."
	Funding: "This work was supported in part by the Department of Mechanical Aerospace and Nuclear Engineering Department at Rensselaer Polytechnic Institute, Troy, NY."

CDC: Centers for Disease Control and Prevention; **COI:** conflicts of interest; **COVID-19:** coronavirus disease 2019; **EU:** European Union; **MERS:** Middle East respiratory syndrome; **NPI:** non-pharmaceutical intervention;**n.r.:** not reported;**PCR:** polymerase chain reaction; **R(T):** current reproduction number; **RT-PCR:** reverse transcription polymerase chain reaction; **SARS:** severe acute respiratory syndrome; **SEIR:** susceptible, exposed, infectious, recovered; **WHO:** World Health Organization

Characteristics of excluded studies *[ordered by study ID]*

Study	**Reason for exclusion**
Adiga 2020	The study does not assess the impact of a travel-related control measure affecting travel across national borders

Study	Reason for exclusion
Aleta 2020	The study does not assess the impact of a travel-related control measure affecting travel across national borders
Annan 2015	The study does not assess the impact of a travel-related control measure affecting travel across national borders
Aravindakshan 2020	The study does not provide a specific effect measure for the travel-related control measure
Arino 2020	The study does not assess the impact of a travel-related control measure affecting travel across national borders
Barkan 2020	The study does not assess the impact of a travel-related control measure affecting travel across national borders
Batista 2020	The study does not assess the impact of a travel-related control measure affecting travel across national borders
Bell 2004	The study does not assess a primary outcome of relevance
Benkouiten 2013	The study does not assess the impact of a travel-related control measure affecting travel across national borders
Benkouiten 2014	The study does not assess the impact of a travel-related control measure affecting travel across national borders
Benkouiten 2015	The study does not assess the impact of a travel-related control measure affecting travel across national borders
Borracci 2020	The study does not provide a specific effect measure for the travel-related control measure
Branas 2020	The study does not provide a specific effect measure for the travel-related control measure
Brauer 2008	The study does not assess COVID-19
Camitz 2006	The study does not assess the impact of a travel-related control measure affecting travel across national borders
Channapathi 2020	The study does not assess the impact of a travel-related control measure affecting travel across national borders
Cowling 2020	The study does not provide a specific effect measure for the travel-related control measure
Dandekar 2020	The study does not provide a specific effect measure for the travel-related control measure
Daon 2020	The study does not assess the impact of a travel-related control measure affecting travel across national borders
Dell'Omodarme 2005	The study does not assess COVID-19
Dursun 2020	The study does not assess a primary outcome of relevance
Eksin 2020	The study does not assess COVID-19
Erandi 2020	The study does not assess the impact of a travel-related control measure affecting travel across national borders

Study	Reason for exclusion
Espinoza 2020	The study does not assess the impact of a travel-related control measure affecting travel across national borders
Fang 2020	The study does not provide a specific effect measure for the travel-related control measure
Fouquet 2020	The study does not assess the impact of a travel-related control measure affecting travel across national borders
Fredj 2020	The study does not assess the impact of a travel-related control measure affecting travel across national borders
Gardner 2016	The study does not assess the impact of a travel-related control measure affecting travel across national borders
Gatto 2020	The study does not assess the impact of a travel-related control measure affecting travel across national borders
Gautret 2013a	The study does not assess the impact of a travel-related control measure affecting travel across national borders
Gautret 2013b	The study does not assess the impact of a travel-related control measure affecting travel across national borders
Gautret 2014	The study does not assess the impact of a travel-related control measure affecting travel across national borders
German 2015	The study does not assess the impact of a travel-related control measure affecting travel across national borders
Gill 2020	The study does not assess the impact of a travel-related control measure affecting travel across national borders
Godin 2021	The study does not assess the impact of a travel-related control measure affecting travel across national borders
Griffiths 2016	The study does not assess the impact of a travel-related control measure affecting travel across national borders
Gunthe 2020	The study does not assess the impact of a travel-related control measure affecting travel across national borders
Hossain 2020	The study does not assess the impact of a travel-related control measure affecting travel across national borders
Hossein 2020	The study does not provide a specific effect measure for the travel-related control measure
Huang 2020	The study did not assess a cross-border impact of a measure related to international travel
Hufnagel 2004	The study does not assess a primary outcome of relevance
Jia 2020	The study does not assess the impact of a travel-related control measure affecting travel across national borders
Johansson 2011	The study does not assess COVID-19
Joo 2019	The study does not assess a primary outcome of relevance

Study	Reason for exclusion
Jungerman 2017	The study does not assess a primary outcome of relevance
Kong 2020	The study does not assess the impact of a travel-related control measure affecting travel across national borders
Kraemer 2020	The study does not assess the impact of a travel-related control measure affecting travel across national borders
Krisztin 2020	The study does not assess the impact of a travel-related control measure affecting travel across national borders
Lai CC 2020	The study did not assess a cross-border impact of a measure related to international travel
Lai S 2020a	The study does not assess the impact of a travel-related control measure affecting travel across national borders
Lai S 2020b	The study does not assess the impact of a travel-related control measure affecting travel across national borders
Lai S 2020c	The study does not assess the impact of a travel-related control measure affecting travel across national borders
Lam 2020	The study does not provide a specific effect measure for the travel-related control measure
Lau 2004	The study does not assess the impact of a travel-related control measure affecting travel across national borders
Lee 2020	The study does not assess a primary outcome of relevance
Li H 2020	The study does not assess the impact of a travel-related control measure affecting travel across national borders
Lin YC 2020	The study does not assess a primary outcome of relevance
Lin YH 2020	No full text is available
Li R 2020	The study does not provide a specific effect measure for the travel-related control measure
Liu 2011	The study does not assess the impact of a travel-related control measure affecting travel across national borders
Liu F 2020	The study does not assess a primary outcome of relevance
Liu Q 2020	The study does not assess COVID-19
Luna 2007	The study does not assess the impact of a travel-related control measure affecting travel across national borders
Ma 2017	The study does not assess the impact of a travel-related control measure affecting travel across national borders
Maeno 2016	The study does not assess the impact of a travel-related control measure affecting travel across national borders
Magalis 2020	The study does not provide a specific effect measure for the travel-related control measure

Study	Reason for exclusion
Malmberg 2020	The study does not assess COVID-19
Marcelino 2012	The study does not assess COVID-19
Myers 2020	The study does not assess the impact of a travel-related control measure affecting travel across national borders
Ng 2020a	The study does not assess the impact of a travel-related control measure affecting travel across national borders
Nikolaou 2020	The study does not assess COVID-19
Niwa 2020	The study does not assess a primary outcome of relevance
Pan 2020	The study does not assess the impact of a travel-related control measure affecting travel across national borders
Pitman 2005	The study does not provide a specific effect measure for the travel-related control measure
Pullano 2020	The study does not assess the impact of a travel-related control measure affecting travel across national borders
Quilty 2020b	The study does not assess the impact of a travel-related control measure affecting travel across national borders
Rajabi 2020	The study does not assess the impact of a travel-related control measure affecting travel across national borders
Ruktanonchai 2020	The study does not assess the impact of a travel-related control measure affecting travel across national borders
Ryu 2019	The study does not assess the impact of a travel-related control measure affecting travel across national borders
Shah 2020	The study does not assess the impact of a travel-related control measure affecting travel across national borders
Shumway 2020	The study does not assess the impact of a travel-related control measure affecting travel across national borders
Sriwijitalai 2020b	The study does not assess the impact of a travel-related control measure affecting travel across national borders
Summan 2020	The study does not assess a primary outcome of relevance
Sun 2013	The study does not assess COVID-19
Thomas 2014	The study does not assess the impact of a travel-related control measure affecting travel across national borders
Valba 2020	The study does not assess COVID-19
Wells 2020b	The study does not assess the impact of a travel-related control measure affecting travel across national borders
Wickramaarachchi 2020	The study does not provide a specific effect measure for the travel-related control measure

Study	Reason for exclusion
Yip 2007	The study does not provide a specific effect measure for the travel-related control measure
Yuan 2020	The study does not assess the impact of a travel-related control measure affecting travel across national borders
Zhao Q 2020	The study does not provide a specific effect measure for the travel-related control measure
Zhao Z 2020	The study does not provide a specific effect measure for the travel-related control measure
Zheng 2020	The study does not assess the impact of a travel-related control measure affecting travel across national borders

COVID-19: coronavirus disease 2019; **MERS:** Middle East respiratory syndrome; **SARS:** severe acute respiratory syndrome

ADDITIONAL TABLES

Table 1. Summary of QUADAS-2 'risk of bias' assessment for screening studies

Study	D1: Traveller selection	D2: Index test	D3: Reference test	D4: Flow and timing
Symptom screening				
Al-Qahtani 2020	Low	Unclear	Low	Unclear
Arima 2020	Unclear	Low	Low	Unclear
Chen J 2020	Low	Unclear	Low	Unclear
Hoehl 2020	Low	Low	Unclear	Unclear
Kim 2020	Low	Low	Unclear	Low
Lytras 2020	Low	Unclear	High	Unclear
Ng 2020	High	Low	Low	Unclear
Wong J 2020	Unclear	Unclear	Unclear	Unclear
Yamahata 2020	Low	Unclear	High	High
PCR test				
Arima 2020	Unclear	Low	Low	Unclear
Al-Qahtani 2020	Low	Low	Low	Unclear
Al-Tawfiq 2020	High	Low	Low	Low
Lagier 2020	High	Unclear	Low	Unclear
Lio 2020	Unclear	Low	Low	Low

Table 1. Summary of QUADAS-2 'risk of bias' assessment for screening studies *(Continued)*

Ng 2020	High	Low	High	Unclear
Shaikh Abdul Karim 2020	Low	Low	Low	Unclear
Combined				
Al-Qahtani 2020	Low	Low	Unclear	Unclear
Arima 2020	Unclear	Low	Low	Unclear
Chen J 2020	Low	Low	Low	Unclear
Lio 2020	Unclear	Low	Low	Low
Shaikh Abdul Karim 2020	Low	Low	Unclear	Unclear

Table 2. Summary of quality appraisal for modelling studies

Study	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10
Travel restrictions reducing or stopping cross-border travel										
Adekunle 2020	Moderate	Moderate	No to minor	Moderate	Reported	No to minor	Not reported	Moderate	No to minor	Moderate
Anderson 2020	No to minor	Moderate	No to minor	Moderate	Reported	No to minor	Not reported	Moderate	Moderate	No to minor
Anzai 2020	Moderate	Major	No to minor	Major	Reported	No to minor	Not reported	Moderate	Moderate	Moderate
Banholzer 2020	No to minor	Major	No to minor	No to minor	Reported	No to minor	Not reported	Moderate	No to minor	No to minor
Binny 2020	No to minor	No to minor	No to minor	No to minor	Not reported	Moderate	Not reported	Moderate	No to minor	Moderate
Boldog 2020	No to minor	No to minor	No to minor	Moderate	Not reported	Moderate	Not reported	Moderate	Moderate	No to minor
Chen T 2020	No to minor	No to minor	No to minor	No to minor	Reported	Moderate	Not reported	Moderate	No to minor	Moderate
Chinazzi 2020	No to minor	No to minor	No to minor	Moderate	Reported	Moderate	Not reported	Moderate	No to minor	Moderate
Costantino 2020	No to minor	Major	No to minor	Moderate	Reported	Moderate	Not reported	Moderate	No to minor	Moderate
Davis 2020	No to minor	No to minor	Moderate	Moderate	Reported	Moderate	Not reported	Moderate	No to minor	Moderate
Deeb 2020	No to minor	No to minor	No to minor	No to minor	Reported	No to minor	Not reported	Moderate	Major	Moderate
Grannell 2020	No to minor	Major	Moderate	No to minor	Not reported	Moderate	Not reported	Moderate	Moderate	Major
Kang 2020	Moderate	Major	No to minor	Major	Reported	No to minor	Not reported	Moderate	Major	Major

Table 2. Summary of quality appraisal for modelling studies *(Continued)*

Liebig 2020	Moderate	Moderate	Moderate	No to minor	Not reported	Moderate	Not reported	Moderate	Major	Major
Linka 2020a	No to minor	Moderate	No to minor	Moderate	Reported	Moderate	Not reported	Moderate	Major	Moderate
Linka 2020b	No to minor	Moderate	No to minor	No to minor	Reported	Moderate	Not reported	Moderate	Major	Moderate
McLure 2020	No to minor	No to minor	No to minor	No to minor	Not reported	Moderate	Not reported	Moderate	Major	No to minor
Nakamura 2020	Moderate	Moderate	No to minor	Moderate	Not reported	Moderate	Not reported	Moderate	Major	Moderate
Nowrasteh 2020	Moderate	Moderate	No to minor	Major	Reported	No to minor	Not reported	Moderate	No to minor	No to minor
Odendaal 2020	Moderate	Major	No to minor	No to minor	Reported	Moderate	Not reported	Moderate	Major	Moderate
Pinotti 2020	No to minor	No to minor	No to minor	No to minor	Reported	No to minor	Not reported	Moderate	Major	Moderate
Russell TW 2020	No to minor	No to minor	No to minor	No to minor	Reported	Moderate	Not reported	Moderate	Moderate	No to minor
Shi 2020	No to minor	Major	No to minor	No to minor	Reported	Moderate	Not reported	Moderate	Major	No to minor
Sruthi 2020	Moderate	Major	No to minor	No to minor	Reported	Moderate	Not reported	Moderate	Moderate	No to minor
Utsunomiya 2020	No to minor	Moderate	No to minor	No to minor	Reported	No to minor	Reported	No to minor	Major	No to minor
Kwok 2020	Moderate	Major	Moderate	Moderate	Not reported	Moderate	Not reported	Moderate	Major	Moderate
Wells 2020	No to minor	No to minor	No to minor	Moderate	Reported	No to minor	Not reported	Moderate	No to minor	No to minor
Yang 2020	No to minor	No to minor	No to minor	No to minor	Reported	Moderate	Not reported	Moderate	No to minor	No to minor

Table 2. Summary of quality appraisal for modelling studies *(Continued)*

Study										
Zhang C 2020	No to minor	Moderate	No to minor	No to minor	Not reported	Moderate	Reported	No to minor	Moderate	No to minor
Zhang L 2020	Moderate	Major	No to minor	No to minor	Reported	Moderate	Not reported	Moderate	Major	Major
Zhong 2020	Moderate	No to minor	No to minor	Moderate	Reported	Moderate	Reported	No to minor	Moderate	Moderate
Screening at borders										
Bays 2020	No to minor	Major	No to minor	Moderate	Not reported	Moderate	Not reported	Moderate	Major	No to minor
Clifford 2020a	No to minor	No to minor	No to minor	Major	Not reported	Moderate	Not reported	Moderate	No to minor	No to minor
Clifford 2020b	No to minor	No to minor	No to minor	No to minor	Not reported	Moderate	Not reported	Moderate	No to minor	No to minor
Dickens 2020	Moderate	Major	No to minor	Moderate	Not reported	Moderate	Not reported	Moderate	Major	Moderate
Gostic 2020	No to minor	Moderate	No to minor	Moderate	Not reported	Moderate	Not reported	Moderate	No to minor	No to minor
Mandal 2020	No to minor	Major	Moderate	Major	Not reported	Moderate	Not reported	Moderate	Moderate	Moderate
Nuckchady 2020	No to minor	Major	No to minor	Major	Reported	Moderate	Not reported	Moderate	Major	Moderate
Quilty 2020	No to minor	Moderate	No to minor	Major	Not reported	Moderate	Not reported	Moderate	Major	No to minor
Russell WA 2020	No to minor	Moderate	Moderate	Moderate	Not reported	Moderate	Not reported	Moderate	No to minor	No to minor
Steyn 2020	No to minor	No to minor	No to minor	No to minor	Not reported	Moderate	Not reported	Moderate	Major	Moderate
Taylor 2020	No to minor	No to minor	No to minor	Moderate	Not reported	Moderate	Not reported	Moderate	Major	No to minor

Table 2. Summary of quality appraisal for modelling studies *(Continued)*

Wells 2020	No to minor	No to minor	No to minor	Moderate	Reported	No to minor	Not reported	Moderate	No to minor	No to minor
Wilson 2020	No to minor	Major	No to minor	Major	Not reported	Moderate	Not reported	Moderate	Major	Moderate
Quarantine of travellers alone										
Ashcroft 2020	No to minor	No to minor	No to minor	No to minor	Not reported	Moderate	Not reported	Moderate	Major	No to minor
Chen Y-H 2020	No to minor	Moderate	No to minor	Moderate	Not reported	Moderate	Not reported	Moderate	Major	Moderate
Chen T 2020	No to minor	No to minor	No to minor	No to minor	Reported	Moderate	Not reported	Moderate	No to minor	Moderate
Clifford 2020b	No to minor	No to minor	No to minor	No to minor	Not reported	Moderate	Not reported	Moderate	No to minor	No to minor
Dickens 2020	Moderate	Major	No to minor	Moderate	Not reported	Moderate	Not reported	Moderate	Major	Moderate
James 2020	No to minor	No to minor	Moderate	No to minor	Not reported	Moderate	Not reported	Moderate	Major	Moderate
Kivuti-Bito 2020	No to minor	No to minor	No to minor	No to minor	Reported	Moderate	Not reported	Moderate	Moderate	Moderate
Russell WA 2020	No to minor	Moderate	Moderate	Moderate	Not reported	Moderate	Not reported	Moderate	No to minor	No to minor
Ryu 2020	No to minor	Major	Moderate	Major	Reported	Moderate	Not reported	Moderate	Moderate	Moderate
Steyn 2020	No to minor	No to minor	No to minor	No to minor	Not reported	Moderate	Not reported	Moderate	Major	Moderate
Taylor 2020	No to minor	No to minor	No to minor	Moderate	Not reported	Moderate	Not reported	Moderate	Major	No to minor
Wong MC 2020	Moderate	Moderate	No to minor	No to minor	Reported	No to minor	Not reported	Moderate	Moderate	Major

Table 2. Summary of quality appraisal for modelling studies (Continued)

Quarantine of travellers and screening combined

Ashcroft 2020	No to minor	No to minor	No to minor	No to minor	Not reported	Moderate	Not reported	Moderate	Major	No to minor
Bays 2020	No to minor	Major	No to minor	Moderate	Not reported	Moderate	Not reported	Moderate	Major	No to minor
Clifford 2020b	No to minor	No to minor	No to minor	No to minor	Not reported	Moderate	Not reported	Moderate	No to minor	No to minor
Russell WA 2020	No to minor	Moderate	Moderate	Moderate	Not reported	Moderate	Not reported	Moderate	No to minor	No to minor
Steyn 2020	No to minor	No to minor	No to minor	No to minor	Not reported	Moderate	Not reported	Moderate	Major	Moderate
Taylor 2020	No to minor	No to minor	No to minor	Moderate	Not reported	Moderate	Not reported	Moderate	Major	No to minor
Wilson 2020	No to minor	Major	No to minor	Major	Not reported	Moderate	Not reported	Moderate	Major	Moderate

![Cochrane Library logo] Cochrane
Library

Trusted evidence.
Informed decisions.
Better health.

Cochrane Database of Systematic Reviews

A P P E N D I C E S

Appendix 1. Review protocol

Travel-related control measures to contain the COVID-19 pandemic: a rapid review (protocol)

Authors

Burns J*[1,2], Movsisyan A*[1,2], Biallas R[1,2], Coenen M[1,2], Emmert-Fees K[3], Geffert K[1,2], Hoffmann S[1,2], Horstick O[4], Pfadenhauer LM[1,2], von Philipsborn P[1,2], Sell K[1,2], Strahwald B[1,2], Stratil JM[1,2], Voss S[1,2], Rehfuess EA[1,2]

* These authors have contributed equally.

[1]Institute for Medical Information Processing, Biometry and Epidemiology; Ludwig-Maximilians-University Munich, Germany
[2]Pettenkofer School of Public Health, Munich, Germany
[3]Institute of Health Economics and Health Care Management, Helmholtz Zentrum München, German Research Center for Environmental Health (GmbH)
[4]Heidelberg Institute of Global Health, Heidelberg University, Germany

Background

The first case of the novel coronavirus disease 2019 (COVID-19) was observed in Wuhan (Hubei province of China) in December 2019. Over the following days and weeks, the disease spread further into China and countries of Asia, including Japan, South Korea, and Thailand (WHO, 2020a). Close to mid-March, cases had been reported in over 100 countries across the globe, and on March 11 2020, the World Health Organization (WHO) declared the outbreak a global pandemic (WHO, 2020b)

COVID-19 is caused by SARS-CoV-2 whose characteristics are similar to those of the coronaviruses causing severe acute respiratory syndrome (SARS-CoV-1/SARS) and Middle East Respiratory Syndrome (MERS-CoV/MERS), respectively. However, in comparison with these viruses, SARS-CoV-2 has higher transmissibility and lower pathogenicity (Fani, Teimoori, & Ghafari, 2020).

Considering lack of existing vaccines to prevent and pharmacological interventions to treat COVID-19, a range of non-pharmacological interventions have been put into place by national and subnational governments to contain and mitigate the spread of the disease. One suite of measures implemented very early during the onset of the COVID-19 pandemic are travel-related control measures. These range from measures, such as border screening to more severe measures, such as complete closure of national borders. Starting from February 2020, most countries and regions in the world have implemented some type of travel-related control measure. As the pandemic develops across the globe, with some countries reporting a beginning second wave of infections (Strzelecki, 2020), it is crucial to understand the effectiveness of these measures to inform decisions on their further (re)implementation, adaptation, relaxation or suspension. This is in line with the International Health Regulations 2005 calling to ground public health decision-making in scientific evidence (WHO, 2005).

Previous reviews have assessed the effectiveness of travel-related control measures, such as international travel bans and entry and/or exit screening in the context of SARS and MERS epidemics (Errett, Sauer, & Rutkow, 2020; Mouchtouri et al., 2019). In May 2020, WHO requested the review authors to develop an evidence map charting the evidence of various travel-related control measures relevant for containing COVID-19 pandemic. The map identified 122 studies assessing a range of measures undertaken across the globe to address COVID-19, SARS, MERS, and influenza. Studies used methods ranging from simple observational approaches to complex modelling techniques. This map informed the scope of the present rapid review to synthesize evidence on the effectiveness of travel-related control measures in coronavirus outbreaks.

Objectives

To inform WHO recommendations and/or national strategies on travel-related control measures to contain the COVID-19 pandemic, this rapid review aims to assess the effectiveness of international travel-related control measures during coronavirus pandemics on infectious disease and screening-related outcomes.

Methods

To conduct this rapid review, at certain stages we will employ abridged procedures of systematic reviewing according to the Cochrane standard (Garritty et al. 2020). Specifically, only one review author will conduct data extraction and risk of bias assessment of epidemiological studies and quality assessment of modelling studies. A second review author will check for correctness, and any uncertainties will be discussed with a third review author. To ensure that this does not compromise the methodological rigor of the systematic review, but also to ensure that all stages of the review are conducted consistently and correctly, we will assign these data extraction and risk of bias/quality assessment tasks to experienced Cochrane review authors and consult researchers with modelling expertise to assist with the quality assessment of modelling studies. Furthermore, we will pilot the procedures for each stage, conduct regular team meetings and keep a list of rolling questions that will be updated continuously.

Criteria for considering studies for this review

Types of studies

In the context of a global pandemic, evidence to inform decisions must be generated rapidly, meaning that methods traditionally used to evaluate impact, such as the randomised controlled trial (RCT), while possible, may not be considered feasible, appropriate, timely or ethical. To ensure that we capture all relevant study types, we will consider a broad range of empirical studies of any size that provide a quantitative measure of impact, including experimental and quasi-experimental studies, observational studies, and epidemiological and mathematical modelling studies.

Experimental and quasi-experimental studies

● Randomized controlled trials (RCTs)

● Interrupted time series (ITS) studies

● Controlled before-after (CBA) studies and difference-in-differences (DiD) studies

● Instrumental variable (IV) studies

● Regression discontinuity (RD) studies

Observational studies

● Cohort studies

● Case-control studies

Mathematical and epidemiological modelling studies

● Compartmental models (e.g. SEIR-type models comprising multiple compartments, such as S: susceptible, E: exposed, I: infectious, R: recovered)

● Bayesian hierarchical models (i.e. models comprising several submodels to integrate observed data as well as uncertainty)

● Spatial models (i.e. modelling disease transmission spatially)

● Epidemiological models (e.g. time series models that model the temporal nature of disease transmission using time-series techniques)

This is likely not an exhaustive list of all the relevant experimental, quasi-experimental, observational, mathematical and epidemiological modelling studies, and - independent of the actual study design - studies use a variety of labels. We will consider all studies providing a quantitative measure of impact, regardless of whether they are indicated by one of these labels.

We will consider studies published in journals as well as those published on preprint servers. We will include RCTs that have been registered but not yet published (in a peer-reviewed journal or on a preprint server) as "ongoing studies".

We will exclude the following types of studies and publications:

● Case reports

● Studies that do not provide a quantitative measure of impact (e.g., studies providing a graphical summary of the development of the number of cases over time in relation to the introduction of control measures, qualitative studies)

● Diagnostic studies (e.g. assessing the sensitivity and specificity of different screening tests)

● Non-empirical studies (e.g. commentaries, editorials, literature reviews not reporting primary empirical data)

● Systematic reviews (although these will be used for backward citation tracking)

● Conference abstracts

Language

Where possible in the compressed time frame, we will consider studies published in all languages. Within the review team, we can consider studies in Armenian, English, French, German, Italian, and Russian. For studies in other languages, we will use existing networks within WHO and Cochrane to support us with screening and/or translation.

Types of participants

We will include human populations (without any age restriction) susceptible to human coronavirus diseases, namely SARS-CoV-1/SARS, SARS-CoV-2/COVID-19, and MERS-CoV/MERS. To be eligible, modelling studies must use modelling parameters for disease transmission specified to reflect one of these diseases.

We will exclude studies (i) not targeting human transmission, (ii) concerned with humans at risk of developing other infectious diseases characterized by different transmission properties (e.g. Ebola and viral meningitis, the transmission mode of which is primarily person-to-person rather than airborne), and/or (iii) addressing humans at risk of developing other infectious diseases, for which travel-related control measures do not play a significant role in containing outbreaks (e.g. influenza).

Types of interventions

We will consider travel-related control measures affecting human travel across national borders. We will consider both introduction/implementation, as well as relaxation of the following measures:

● Closure of national borders to entry and/or exit

● International travel restrictions/bans

○ Denial of entry and/or exit on the basis of nationality, travel history, health status or other characteristics

○ Full or partial suspension of cross-border travel via land and/or air and/or sea

○ Visa requirement or refusal on the basis of nationality, travel history, health status or other characteristics

● Entry and/or exit screening at national borders

○ Temperature measurement

○ Health questionnaire (e.g. symptoms, travel history, contact tracing)

○ Thermography

○ Physical examination

○ Testing for current or past infection

○ Passive observation

● Quarantine or isolation of travellers crossing national borders

● Any combination of the above measures

We will exclude the following types of interventions:

● Combinations of the above-mentioned travel-related control measures with other measures where studies do *not* provide effect estimates for the travel-related control measures (e.g., studies providing a combined effect estimate for suspension of cross border travel and use of mandatory face masks). Studies, where the effect of travel-related control measures cannot be disentangled from the effect of a broader suite of public health measures, cannot usefully inform WHO recommendations on whether countries should or should not consider travel-related control measures to contain the COVID-19 pandemic (see review objective).

● All interventions not directly related to travel, including a range of containment and mitigation measures (e.g., community-based quarantine, personal protective measures, hygiene measures, bans on mass gatherings and other social-distancing measures)

● All interventions related to movement of animals or goods

● All interventions concerned with human travel across subnational borders. While subnational measures can potentially inform national travel-related control measures, these measures are not prioritised by the WHO, and as shown in the previous evidence map, are often impossible to disentangle from other subnational measures, such as lockdowns, community quarantine or social distancing recommendations.

● Travel warnings or travel advice issued by the World Health Organization or national governments

● Interventions solely concerned with the accuracy of tests rather than their implementation as part of an entry and/or exit border control measure

● Usual practice (e.g. seasonal changes to travel) or events (e.g. school holidays) affecting travel but not representing travel-related control measures

● Cancellation of events affecting international travel but undertaken as a means to prevent mass gatherings (e.g., Hajj, international sporting events, international trade fairs)

There are two Cochrane rapid reviews that may identify overlapping studies. One published review, which is currently being updated, focuses on quarantine measures, including quarantining travellers crossing national borders. The other review, currently under editorial assessment and soon to be published, focuses on screening measures, including entry and/or exit screening at national borders. In discussions with Cochrane and WHO, we decided that it would be important for decision-makers to be able to access the evidence on all travel-related control measures in a single review. To address overlap between the present review and the two separately conducted reviews, we will coordinate the synthesis and write-up of our review findings with those of the quarantine and screening reviews to ensure consistency in reporting and interpretation of the overlapping questions and evidence.

Types of outcomes

Primary outcomes:

We will consider studies assessing any of the following infectious disease transmission and screening-related outcomes:

● Cases avoided due to the intervention: e.g. number, proportion, rate of cases observed or predicted with and without the intervention

● Number of cases detected due to the intervention: e.g. case detection rate (i.e. number of cases detected per 10,000 screened), positive predictive value (i.e. number of cases detected per those identified as high risk)

● Shift in epidemic development due to the intervention: e.g. probability of epidemic, time to/delay in epidemic arrival or peak, size of epidemic peak, change in the effective reproduction number

Secondary outcomes:

We will consider the following secondary outcomes if identified in studies that assessed at least one of the primary outcomes:

● Any other infectious disease transmission outcomes

● Healthcare utilization (e.g. number of cases requiring ICU treatment, time until ICU capacity is reached)

● Resource requirements for implementing the intervention (e.g. costs associated with intervention, additional personnel, number of tests required)

● Any adverse effects (e.g. health, economic, and social outcomes)

We will not consider studies reporting on other outcomes.

Search methods for identification of studies

The search strategy will be structured around two blocks focusing on (1) COVID-19, SARS and MERS and (2) travel-related control measures. We will conduct the searches in English but will aim to include studies published in any language (see "Language" above). We will adapt the search strategy used in the evidence map for travel-related control measures. An experienced information specialist will adapt and run systematic searches in the following electronic databases up to the last week of June, 2020:

● Ovid MEDLINE(R) and Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Daily and Versions(R) (1946-present)

● Ovid Embase (1996-present)

We will additionally search the following COVID-19-specific databases:

● Cochrane COVID-19 Register (https://covid-19.cochrane.org/) ,which contain study references from ClinicalTrials.gov, WHO International Clinical Trials Registry Platform (ICTRP), PubMed, medRxiv and other hand-search articles from publishers' websites.

● WHO COVID-19 Global literature on coronavirus disease (https://search.bvsalud.org/global-literature-on-novel-coronavirus-2019-ncov/), which contains primarily research (published AND/OR prepublication) journal articles from PubMed, Web of Science, Global Index Medicus, Embase. In addition, Lanzhou University submits on a daily basis citations from CNKI as well as a number of Chinese journal publishers.

● CDC COVID-19 Research Articles Downloadable Database (https://www.cdc.gov/library/researchguides/2019novelcoronavirus/researcharticles.html) will be searched exclusively for preprints from bioRxiv, medRxiv and SSRN.

Finally, we will screen the reference lists of the existing systematic reviews on travel-related control measures to identify additional eligible studies. See Appendix 1 for the MEDLINE search strategy.

Data collection and analysis

Selection of studies

We will develop a standardized title and abstract screening guidance and pilot it among all review authors involved with the screening using the same 100 titles and abstracts. We will discuss and resolve all issues and revise the screening guidance accordingly. One review author will then screen all titles and abstracts. A second review author will screen all excluded abstracts. We will adopt an inclusive approach at this stage, and all unclear studies will be taken forward to the full text screening. Screening all full texts in duplicate, as described below, will ensure that the single screening of titles and abstracts included at this stage will not compromise methodological rigor.

We will conduct a pilot of the full text screening; all review authors involved with full text screening will screen a set of ten studies at the outset of this stage (Garritty et al., 2020). The team will then discuss any open questions or issues, as well as how to harmonize screening across all review authors. Following the pilot, two review authors will screen the remaining studies in duplicate. Any discrepancies between the two review authors will be discussed, and a third review author or the entire author team will be consulted where necessary.

Data extraction and management

One review author will extract study characteristics and data from all included studies using a data extraction form in Microsoft Excel. All extracted data will be checked by a second review author. The data extraction form will be an extended version of the form used in the previous evidence map. The following categories will be included in the extraction form:

● Study type

● Population, setting and context

● Characteristics of pathogen/disease

● Context of pathogen/disease

● Intervention

● Outcomes and results

We will pilot the data extraction form using five studies meeting the inclusion criteria. Appendix 2 provides the details on the data extraction categories.

Based on our experience with the evidence map on travel related-control measures, we expect key aspects, notably the description of the intervention and the stage of the epidemic at which the intervention was implemented, to be poorly reported in included studies. Therefore, for included studies addressing COVID-19 in which data on the intervention and stage of the pandemic are not well reported, and where applicable (e.g., some modelling studies may not directly relate to a specific country or point in time), we will explore whether searching for and extracting additional data from external sources will enhance our understanding of these aspects. To do so, we will consider data from the following external sources:

● Intervention data: the COVID-19 Country Policy Tracker, offered by the Organisation for Economic Cooperation and Development (OEC), collates and summarizes countries' policy responses to the current pandemic (https://www.oecd.org/coronavirus/country-policy-tracker/). ACAPS, an independent information provider, also manages a large database of country-level policy responses (https://www.acaps.org/covid19-government-measures-dataset).

● Stage of the pandemic: Daily COVID-19 situation reports published by the WHO summarize the daily number of total and new cases and deaths for all countries, and specify the current stage of the pandemic as "sporadic cases", "clusters of cases" or "community transmission" (https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports/).

In cases for which we utilize external sources, we will carefully document the source of the information.

Assessment of risk of bias in included studies

One review author will rate the risk of bias (RoB) or quality of each included study, depending on the type of study, and a second review author will check these ratings; any questions or uncertainties that arise throughout the process will be discussed between these review authors or among the review team.

Given the broad range of study designs, we will apply multiple tools in assessing RoB of included studies. For any identified experimental studies, we will apply the Cochrane RoB 2 tool (Higgins et al., 2019). For quasi-experimental and observational studies, we will apply ROBINS-I (Sterne et al., 2016). In applying ROBINS-I, important confounding factors that each study would ideally control for should be defined a priori. Based on the body of evidence identified in the recently conducted evidence map, we expect that most non-modelling studies will assess the impact of entry and/or exit screening at national borders on the number of cases detected. For those studies, we outline two important categories of confounding factors:

· Between-group differences: where multiple groups are assessed (i.e., screened and non-screened), these regard the composition of the two groups with respect to age, sex, socioeconomic status and other factors potentially affecting the risk of infection (e.g., occupation, stay in a high-prevalence area);

· Temporal changes: as these studies are longitudinal in nature, changes over time could lead to confounding (i.e. changes to testing practices, training provided to personnel, and changes in implementation of the measures).

No validated tool exists for assessing the RoB of modelling studies, and little consensus exists in the systematic review community about how to best approach the assessment of RoB and/or quality of an individual modelling study. A rapid review of the methodological literature aimed to identify and summarize studies describing criteria for assessing the quality of mathematical studies (Egger et al. 2017). This review identified 20 studies (including Jaime Caro et al., 2014; Philips, Bojke, Sculpher, Claxton, & Golder, 2006) that varied in scope and level of detail regarding the proposed criteria; looking across the identified studies, it suggested that an assessment of the quality of a modelling study should capture the aspects of (i) model structure, (ii) input data, (iii) different dimensions of uncertainty, (iv) transparency, and (v) validation. Following the suggestions by Eggers et al. (2017), we developed a bespoke tool for the assessment of modelling studies and selected single criteria from two studies (Jaime Caro et al., 2014; Philips, Bojke, Sculpher, Claxton, & Golder, 2006). We sought to strike a balance between capturing all five key aspects and feasibility of applying the criteria. The individual criteria we will apply are outlined in Appendix 3. We will report each of the criteria separately, i.e. we will not combine multiple criteria into a summary score. The quality of modelling studies will be assessed by one review author and checked by a second.

Data synthesis

We will synthesize the findings narratively, graphically, or in tabular form, stratified by intervention type and disease. Drawing on the previous evidence map and the large heterogeneity identified with regard to the setting, population, intervention and other contextual factors, we expect that data will not be sufficiently similar to conduct meta-analyses. Where possible, we will summarize effect estimates graphically using forest plots.

Subgroup analyses and investigation of heterogeneity

In the absence of meta-analyses, we will not conduct a statistical assessment of heterogeneity, nor will we statistically assess differences between subgroups. We will nevertheless investigate the influence of potentially important sources of heterogeneity on the impact of interventions, focusing on heterogeneity in selected PICO elements and context. We will consider the following sources of heterogeneity:

● Stage of the pandemic during which the intervention was implemented: impact of travel-related control measures are likely different in countries where only few cases have been detected than in countries where community transmission is established;

● Whether a country is an island state or not: the geographical characteristics of the state (i.e., island state versus countries with land borders) will serve as a proxy measure of the interconnectedness of the countries and explain possible differences in the efficiency in managing and controlling entries/exits at all ports of entry;

● Assumed infectious disease parameters: assumptions used in models may have a large influence on the direction and magnitude of reported results;

● Whether a single travel-related control measure or a package of travel-related control measures were implemented: multiple integrated travel-related interventions may be more effective than single stand-alone interventions.

We will explore heterogeneity narratively; forest plots and/or tables will be stratified according to these specific aspects. This will not allow us to derive conclusions but could point to aspects potentially modifying the effectiveness of different measures.

Sensitivity analyses

Where a sufficiently homogeneous evidence base is identified, we will conduct sensitivity analyses, excluding experimental, quasi-experimental, or observational studies at high RoB or modelling studies with major quality concerns, to assess whether these studies influence the results. As for the subgroup analyses described above, in the absence of meta-analyses, this assessment will be narrative, not statistical, in nature.

Assessment of certainty of evidence

We will use the GRADE approach to assess the certainty of the primary outcomes. One review author will collate the evidence per primary outcome and suggest initial certainty of evidence ratings. These will then be further deliberated in a team of review authors and a joint decision for certainty of evidence ratings will be made for each primary outcome.

The certainty of evidence is defined in GRADE as the extent to which one can be confident that the true effect of an intervention lies on one side of a specified threshold, or within a chosen range (Hultcrantz et al., 2017). In this rapid review, we will consider "difference from the null" as an important threshold assuming that even the small effect sizes may be relevant for population-level travel-related control measures.

The certainty of evidence rating in GRADE yields four possible levels of evidence: high certainty (i.e., the estimated effect lies close to the true effect), moderate certainty (i.e., the estimated effect is probably close to the true effect), low certainty (i.e., the estimated effect might substantially differ from the true effect), and very low certainty (i.e., the estimated effect is probably substantially different from the true effect).

We will rate bodies of evidence from experimental, observational and modelling studies separately. In GRADE, evidence from RCTs enters the rating as high certainty, as does evidence from observational studies whose risk of bias has been assessed using ROBINS-I (Schunemann et al., 2019). Further to this five domains are used to further downgrade evidence, including *study limitations, inconsistency, indirectness, imprecision* and *publication bias*, and three domains are used to upgrade evidence, including *plausible confounding, large estimates of effect,* and *dose-response relationship*. These domains apply to assessment of evidence from all types of studies, including modelling studies.

To rate certainty of evidence from modelling studies, we will use the recent guidance developed by the GRADE Working Group (Brozek et al.). Evidence from modelling studies also starts the assessment as high certainty, and all the GRADE domains described above are used to assess certainty of model outputs.

Disclaimer: This rapid review is being commissioned and paid for by the World Health Organization (WHO). The authors alone are responsible for the views expressed in this protocol and they do not necessarily represent the views, decisions or policies of WHO.

References

Brozek, J., Canelo-Aybar, C., Akl, E. A., Bowen, J. M., Butcher, J., Chiu, W. A., & Cronin, M. GRADE Guidelines 30: the GRADE approach to assessing the certainty of modelled evidence - an overview in the context of health decision-making. (Draft article).

Egger, M., Johnson, L., Althaus, C., Schoni, A., Salanti, G., Low, N., & Norris, S. L. (2017). Developing WHO guidelines: Time to formally include evidence from mathematical modelling studies. *F1000Res, 6*, 1584.

Errett, N. A., Sauer, L. M., & Rutkow, L. (2020). An integrative review of the limited evidence on international travel bans as an emerging infectious disease disaster control measure. *Journal of emergency management (Weston, Mass.), 18*(1), 7-14.

Fani, M., Teimoori, A., & Ghafari, S. (2020). Comparison of the COVID-2019 (SARS-CoV-2) pathogenesis with SARS-CoV and MERS-CoV infections. *Future Virol.* (Epub ahead of print).

Garritty, C., Gartlehner, G., Kamel, C., King, V. J., Nussbaumer-Streit, B., Stevens, A., . . . Affengruber, L. (2020). Cochrane Rapid Reviews. Interim Guidance from the Cochrane Rapid Reviews Methods Group.

Higgins, J. P. T., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M. J., & Welch, V. A. e. (2019). *Cochrane handbook for systematic reviews of interventions.* (2nd ed.). Chichester (UK): John Wiley & Sons.

Hultcrantz, M., Rind, D., Akl, E. A., Treweek, S., Mustafa, R. A., Iorio, A., . . . Guyatt, G. (2017). The GRADE Working Group clarifies the construct of certainty of evidence. *J Clin Epidemiol, 87*, 4-13.

Jaime Caro, J., Eddy, D. M., Kan, H., Kaltz, C., Patel, B., Eldessouki, R., . . . Forces, I.-A.-N. M. C. T. (2014). Questionnaire to assess relevance and credibility of modelling studies for informing health care decision making: an ISPOR-AMCP-NPC Good Practice Task Force report. *Value Health, 17*(2), 174-182.

Mouchtouri, V. A., Christoforidou, E. P., An der Heiden, M., Menel Lemos, C., Fanos, M., Rexroth, U., . . . Hadjichristodoulou, C. (2019). Exit and Entry Screening Practices for Infectious Diseases among Travelers at Points of Entry: Looking for Evidence on Public Health Impact. *International journal of environmental research and public health, 16*(23), 4638-4638.

Nussbaumer-Streit, B., Mayr, V., Dobrescu, A. I., Chapman, A., Persad, E., Klerings, I., . . . Gartlehner, G. (2020). Quarantine alone or in combination with other public health measures to control COVID-19: a rapid review. *Cochrane Database of Systematic Reviews*(4).

Philips, Z., Bojke, L., Sculpher, M., Claxton, K., & Golder, S. (2006). Good practice guidelines for decision-analytic modelling in health technology assessment: a review and consolidation of quality assessment. *Pharmacoeconomics, 24*(4), 355-371.

Schunemann, H. J., Cuello, C., Akl, E. A., Mustafa, R. A., Meerpohl, J. J., Thayer, K., . . . Group, G. W. (2019). GRADE guidelines: 18. How ROBINS-I and other tools to assess risk of bias in nonrandomized studies should be used to rate the certainty of a body of evidence. *J Clin Epidemiol, 111*, 105-114.

Sterne, J. A., Hernan, M. A., Reeves, B. C., Savovic, J., Berkman, N. D., Viswanathan, M., . . . Higgins, J. P. (2016). ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ, 355*, i4919.

Strzelecki, A. (2020). The second worldwide wave of interest in coronavirus since the COVID-19 outbreaks in South Korea, Italy and Iran: A Google Trends study. *Brain Behav Immun*. (Article in press).

WHO. (2005). International Health Regulations. WHA 58.3, 2nd ed. Geneva: World Health Organization.

WHO. (2020a). Novel Coronavirus (2019-nCoV) Situation Report - 1. World Health Organization.

WHO. (2020b). Novel Coronavirus (2019-nCoV) Situation Report - 51. World Health Organization.

Protocol Appendix 1. Search Strategy

Database: Ovid MEDLINE(R) and Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Daily and Versions(R) 1946 to June 25, 2020

Date search conducted: June 26, 2020

Strategy:

1 exp Coronavirus/ (18064)

2 Coronavirus Infections/ (12378)

3 COVID-19.rs. (7963)

4 severe acute respiratory syndrome coronavirus 2.os. (6648)

5 (2019 nCoV or 2019nCoV or 2019-novel CoV).ti,ab,kf. (891)

6 (Coronavir* or corona virus* or Middle East Respiratory Syndrome* or MERS or Severe Acute Respiratory Syndrome* or SARS*).ti,ab,kf. (35015)

7 COVID 19.mp. (25238)

8 (COVID19 or COVID 2019).ti,ab,kf. (437)

9 (nCov 2019 or nCov 19).ti,ab,kf. (55)

10 or/1-9 [Set 1: Coronaviruses] (51693)

11 Air Travel/ (366)

12 Travel/ (24964)

13 (border? adj3 (clos* or restrict* or control* or measure?)).ab,kf. (1046)

14 ((isolat* or quarantin*) adj6 (exposed or suspected or travel* or airport? or border?)).ti,ab,kf. (7939)

15 ((mobility or movement*) adj2 (reduc* or restrict*)).ti,ab,kf. (9614)

16 (travel* or border?).ti. (26180)

17 (travel adj4 (measure? or intervention? or NPI?)).ab,kf. (424)

18 (travel* adj3 (restrict* or reduc* or control* or limit* or lockdown? or ban*)).ab,kf. (1545)

19 ((questionnaire* or screen* or surveil*) adj4 (traveller? or entr* or exit or border? or airport?)).ti,ab,kf. (1840)

20 visa?.ti,ab,kf. (2086)

21 or/11-20 [Set 2: Travel measures] (64508)

22 and/10,21 [Sets 1 & 2] (945)

23 limit 22 to "humans only (removes records about animals)" (925)

24 remove duplicates from 23 (916)

Protocol Appendix 2. Categories in the data extraction form

NOTE: aspects in bold will be primarily relevant for modelling studies.

Study information:

- Study ID
- Study title
- Study source (i.e., journal, report, preprint publication)
- Date of submission
- Date of publication

Study type:

- Study type (e.g., experimental, quasi-experimental, observational, modelling)
- Verbal summary of study type (e.g., stochastic discrete event simulation model)
- Comments

Population, setting, and context:

- Country in which travel-related control measure is implemented
- Region protected by travel-related control measure
- Region restricted by travel-related control measure
- Short description of population studied (e.g., international travellers including flight passenger and crew members arriving between 27 April and 22 June 2009)
- Comments

Characteristics of pathogen/disease:

- Disease (i.e., SARS, MERS, COVID-19)
- **Incubation period**
- **Latent period**
- **Infectiousness of symptomatic persons**
- **Infectiousness of asymptomatic persons**
- Comments

Context of pathogen/disease transmission:

- Stage of pandemic in the protected and restricted regions during which the intervention was implemented (i.e., no cases, sporadic cases, clusters of cases, and community transmission)
  * Supporting information for the categorization of the stage of pandemic in the protected and restricted regions (e.g., date of notification of first case, current number of cases, established community spread)
- **Reproduction number: (R0) and R(t) for the region protected by the intervention**

Intervention:

- Broad measure category (i.e., closure of national borders to entry and/or exit, international travel restrictions/bans, entry and/or exit screening at national borders, quarantine or isolation of travelers crossing national borders)
- Verbal summary of specific travel-related measure(s) (e.g., all arriving passengers were provided with a symptom questionnaire)
- Date(s) of implementation of the travel-related control measure(s)
- Any reported exceptions to the measure (e.g., certain individuals being excluded or fast-tracked in airport screening because of nationality, occupation, country of origin)
- **Representation of the intervention in the model (e.g., 100% travel reduction)**
- Comments

Primary outcomes (repeated for each primary outcome):

- Outcome category (i.e., cases avoided due to the intervention, number of cases detected due to the intervention, shift in epidemic development due to the intervention)
- Description of outcome
- Length of follow-up
- Estimate related to the impact of the travel-related control measure
- Narrative summary of overall impact of travel-related control measure
- **Additional analysis used to quantify the range of potential effects for this outcome**
- Comments

Secondary outcomes(repeated for each secondary outcome):

- Description of any secondary (non-quantitative) outcomes (i.e., any other infectious disease transmission outcomes, healthcare utilization outcomes, resource requirements for implementing the intervention, and any adverse effects)
- Narrative results of secondary outcomes
- Comments

[External sources] Stage of pandemic in protected and restricted regions during which the intervention was implemented (i.e., no cases, sporadic cases, clusters of cases, and community transmission) – for each aspect extracted the exact source will be documented

[External sources] Intervention – for each aspect extracted the exact source will be documented

- Verbal summary of specific travel-related measure(s) (e.g., all arriving passengers were provided with a symptom questionnaire)
- Date(s) of implementation of the travel-related control measure(s)
- Any reported exceptions to the measure (e.g., certain individuals being excluded or fast-tracked in airport screening because of nationality, occupation, country of origin)

Protocol Appendix 3. Criteria used for assessing the quality of individual modelling studies

Aspect	Source	Questions
Model structure	Philips et al. 2006	1. Are the structural assumptions transparent and justified?
		2. Are the structural assumptions reasonable given the overall objective, perspective and scope of the model?
Input data	Caro et al. 2014	3. All things considered, do you agree with the values used for the inputs?
Validation (external)	Caro et al. 2014	4. Has the model been shown to accurately reproduce what was observed in the data used to create the model?
		5. Has the model been shown to accurately estimate what actually happened in one or more separate studies?
		6. Has the model been shown to accurately forecast what eventually happens in reality?
Validation (internal)	Caro et al. 2014	7. Have the process of internal verification and its results been documented in detail?
		8. Has the testing been performed systematically?
		9. Does the testing indicate that all the equations are consistent with their data sources?
Different dimensions of uncertainty	Caro et al. 2014	10. Was there an adequate assessment of the effects of uncertainty?
Transparency	Caro et al. 2014	11. Was technical documentation, in sufficient detail to allow (potentially) for replication, made available openly or under agreements that protect intellectual property?

Appendix 2. List of systematic reviews considered for backward citation searches

1. Bielecki M, Patel D, Hinkelbein J, Komorowski M, Kester J, Ebrahim S, et al (2021). Air travel and COVID-19 prevention in the pandemic and peri-pandemic period: a narrative review. Travel Medicine and Infectious Disease, 39.
2. Bitar D, Goubar A, Desenclos JC. (2009). International travels and fever screening during epidemics: a literature review on the effectiveness and potential use of non-contact infrared thermometers. Eurosurveillance, 14(6), 854-8. Retrieved from http://europepmc.org/abstract/MED/19215720. (Accession No. 19215720).

3. Browne A, Ahmad SS, Beck CR, Nguyen-Van-Tam JS (2016). The roles of transportation and transportation hubs in the propagation of influenza and coronaviruses: a systematic review. Journal of Travel Medicine, 23(1). doi:10.1093/jtm/tav002.

4. Burns J, et al (2021). Travel-related control measures to contain the COVID-19 pandemic: an evidence map. BMJ Open [accepted manuscript].

5. Chetty T, Daniels BB, Ngandu NK, Goga A. (2020). A rapid review of the effectiveness of screening practices at airports, land borders and ports to reduce the transmission of respiratory infectious diseases such as COVID-19. South African Medical Journal, 110(11), 1105-1109. doi:10.7196/SAMJ.2020.v110i11.15124.

6. Errett NA, Saue LM, Rutkow L. (2020). An integrative review of the limited evidence on international travel bans as an emerging infectious disease disaster control measure. Journal of Emergency Management, 18(1), 7-14. doi:10.5055/jem.2020.0446.

7. Gautret P, Benkouiten S, Al-Tawfiq JA, Memish ZA. (2016). Hajj-associated viral respiratory infections: A systematic review. Travel Medicine and Infectious Disease, 14(2), 92-109. doi:10.1016/j.tmaid.2015.12.008.

8. Getaneh Y, Yizengaw A, Adane S, Zealiyas K, Abate Z, Leulseged S, et al. (2020). Global lessons and Potential strategies in combating COVID-19 pandemic in Ethiopia: Systematic Review. medRxiv, 2020.2005.2023.20111062. doi:10.1101/2020.05.23.20111062.

9. Imai N, Gaythorpe AM, Abbott S, Bhatia S, van Elsland S, Prem K, et al. (2020). Adoption and impact of non-pharmaceutical interventions for COVID-19 [version 1; peer review: awaiting peer review].

10. Jefferson T, Del Mar CB, Dooley L, Ferroni E, Al-Ansary LA, Bawazeer GA, . . . Conly JM. (2011). Physical interventions to interrupt or reduce the spread of respiratory viruses. Cochrane Database of Systematic Reviews, 2011(7), Cd006207. doi:10.1002/14651858.CD006207.pub4.

11. Kang S, Moon J, Kang H, Nam H, Tak S, Cho S. (2020). The Evolving Policy Debate on Border Closure in Korea. Journal of Preventive Medicine and Public Health 53(5), 302-306. doi:10.3961/jpmph.20.213

12. Lahiri A, Jha SS, Bhattacharya S, Ray S, Chakraborty A. (2020). Effectiveness of preventive measures against COVID-19: A systematic review of In Silico modeling studies in Indian context. Indian Journal of Public Health, 64(Supplement), S156-s167. doi:10.4103/ijph.IJPH_464_20.

13. Mouchtouri VA, Christoforidou EP, An der Heiden M, Menel Lemos C, Fanos M, Rexroth U, et al (2019). Exit and Entry Screening Practices for Infectious Diseases among Travelers at Points of Entry: Looking for Evidence on Public Health Impact. International Journal of Environmental Research and Public Health, 16(23). doi:10.3390/ijerph16234638.

14. Nussbaumer-Streit B, Mayr V, Dobrescu AI, Chapman A, Persad E, Klerings I, et al (2020). Quarantine alone or in combination with other public health measures to control COVID-19: a rapid review. Cochrane Database of Systematic Reviews(9). doi:10.1002/14651858.CD013574.pub2.

15. Patino-Lugo DF, Velez M, Velasquez Salazar P, Vera-Giraldo CY, Velez V, Marin IC, et al. (2020). Non-pharmaceutical interventions for containment, mitigation and suppression of COVID-19 infection. Colombia medica (Cali, Colombia), 51(2), e4266. doi:10.25100/cm.v51i2.4266.

16. Shah SA, Mansor J, Nurumal SR, Wan Ibadullah WAH, Mohammad Z, Rosli NM, Singh PJ. (2020). Rapid response and public health measures of COVID-19 infection among Asian countries. Gazi Medical Journal, 31(2 A), 239-243. doi:10.12996/gmj.2020.63.

17. Tabari P, Amini M, Moghadami M, Moosavi M. (2020). International public health responses to COVID-19 outbreak: A rapid review. Iranian Journal of Medical Sciences, 45(3), 157-169. doi:10.30476/ijms.2020.85810.1537.

18. Viswanathan M, Kahwati L, Jahn B, Giger K, Dobrescu AI, Hill C, et al (2020). Universal screening for SARS-CoV-2 infection: a rapid review. Cochrane Database of Systematic Reviews(9). doi:10.1002/14651858.CD013718.

Appendix 3. Search strategy

Search sources

Database: Ovid MEDLINE(R) and Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Daily and Versions(R) 1946 to 12 November 2020

Date search conducted: 13 November 2020

Strategy:

1   exp Coronavirus/ (41876)

2   Coronavirus Infections/ (40179)

3   COVID-19.rs. (35243)

4   severe acute respiratory syndrome coronavirus 2.os. (29910)

5   (2019 nCoV or 2019nCoV or 2019-novel CoV).ti,ab,kf. (1285)

6   (Coronavir* or corona virus* or Middle East Respiratory Syndrome* or MERS or Severe Acute Respiratory Syndrome* or SARS*).ti,ab,kf. (60294)

7   COVID 19.mp. (68909)

8   (COVID19 or COVID 2019).ti,ab,kf. (1036)

9   (nCov 2019 or nCov 19).ti,ab,kf. (94)

10   or/1-9 [Set 1: Coronaviruses] (97490)

11   Air Travel/ (396)

12   Travel/ (25483)

13   (border? adj3 (clos* or restrict* or control* or measure?)).ab,kf. (1139)

14   ((isolat* or quarantin*) adj6 (exposed or suspected or travel* or airport? or border?)).ti,ab,kf. (8173)

15   ((mobility or movement*) adj2 (reduc* or restrict*)).ti,ab,kf. (10041)

16   ((questionnaire* or RT-PCR or screen* or surveil* or test* or telethermographic* or temperature or thermal imag* or thermal scan* or thermomet* or thermograph*) adj4 (traveller? or entr* or exit or border? or airport?)).ti,ab,kf. (5328)

17   (travel* or border?).ti. (26695)

18   (travel adj4 (measure? or intervention? or NPI?)).ab,kf. (463)

19   (travel* adj3 (restrict* or reduc* or control* or limit* or lockdown? or ban*)).ab,kf. (1806)

20   visa?.ti,ab,kf. (2175)

21   or/11-20 [Set 2: Travel measures] (69713)

22   and/10,21 [Sets 1 & 2] (1700)

23   limit 22 to "humans only (removes records about animals)" (1675)

24   remove duplicates from 23 (1669)

Database: Ovid Embase 1996 to 2020 November 12

Date search conducted: 13 November 2020

Strategy:

1   coronaviridae/ (1048)

2   exp coronavirinae/ (21375)

3   exp coronavirus infection/ (22841)

4   (2019 nCoV or 2019nCoV or 2019-novel CoV).ti,ab,kw. (1294)

5   (Coronavir* or corona virus* or Middle East Respiratory Syndrome* or MERS or Severe Acute Respiratory Syndrome* or SARS*).ti,ab,kw. (62133)

6   COVID 19.af. (65383)

7   (COVID19 or COVID 2019).ti,ab,kw. (1064)

8   (nCov 2019 or nCov 19).ti,ab,kw. (66)

9   or/1-8 [Set 1: Coronaviruses] (103420)

10   air transportation/ (117)

11   aviation/ (7501)

12   travel/ (47727)

13   (border? adj3 (clos* or restrict* or control* or measure?)).ab,kw. (1329)

14   ((isolat* or quarantin*) adj6 (exposed or suspected or travel* or airport? or border?)).ti,ab,kw. (9807)

15   ((mobility or movement*) adj2 (reduc* or restrict*)).ti,ab,kw. (13481)

16   ((questionnaire* or RT-PCR or screen* or surveil* or test* or telethermographic* or temperature or thermal imag* or thermal scan* or thermomet* or thermograph*) adj4 (traveller? or entr* or exit or border? or airport?)).ti,ab,kw. (6673)

17   (travel* or border?).ti. (29615)

18   (travel adj4 (measure? or intervention? or NPI?)).ab,kw. (553)

19   (travel* adj3 (restrict* or reduc* or control* or limit* or lockdown? or ban*)).ab,kw. (2244)

20   visa?.ti,ab,kw. (2493)

21   or/10-20 [Set 2: Travel measures] (105738)

22   and/9,21 [Sets 1 & 2] (2420)

23   (animal experiment/ or exp animal/) not exp human/ (5031646)

24   22 not 23 (2389)

25   conference abstract.pt. (3904744)

26   24 not 25 (2248)

27   remove duplicates from 26 (2204)

Database: Cochrane COVID-19 Study Register

URL: covid-19.cochrane.org (searched via the Cochrane Register of Studies: crsweb.cochrane.org)

Date search conducted: 13 November 2020

Strategy:

#1. (border* OR travel*):TI (102)

#2. (border* ADJ3 (clos* or restrict* or control* or measure*)):TI,AB (37)

#3. ((isolat* or quarantin*) AND (exposed or suspected or travel* or airport* or border*)):TI,AB (564)

#4. ((mobility or movement*) ADJ2 (reduc* or restrict*)):TI,AB (144)

#5. ((questionnaire* or RT-PCR or screen* or surveil* or test* or telethermographic* or temperature or thermal imag* or thermal scan* or thermomet* or thermograph*) AND (traveller* or entr* or exit or border* or airport*)):TI,AB (266)

#6. (travel ADJ4 (measure* or intervention* or NPI*)):TI,AB (37)

#7. (travel* ADJ3 (restrict* or reduc* or control* or limit* or lockdown* or ban*)):TI,AB (162)

#8. visa*:TI,AB (1)

#9. #1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8 (1064)

Database: WHO COVID-19 Global literature on coronavirus disease

URL: search.bvsalud.org/global-literature-on-novel-coronavirus-2019-ncov

Date search conducted: 13 November 2020

Strategy:

(ti:(border OR borders OR travel*)) OR (tw:(border* AND (clos* OR restrict* OR control* OR measure*))) OR (tw:((isolat* OR quarantin*) AND (exposed OR suspected OR travel* OR airport* OR border*))) OR (tw:((mobility OR movement*) AND (reduc* OR restrict*) AND travel*)) OR (tw:((questionnaire* or "RT-PCR" or screen* or surveil* or test* or telethermographic* or temperature or "thermal image" or "thermal images" or "thermal imaging" or "thermal scan" or "thermal scans" or "thermal scanning" or thermomet* or thermograph*) AND (traveller*

OR entr* OR exit OR border* OR airport*))) OR (tw:(travel AND (measure* OR intervention* OR NPI*))) OR (tw:(travel* AND (restrict* OR reduc* OR control* OR limit* OR lockdown* OR ban*))) OR (tw:(visa OR visas)) (5010)

Filters applied: WHO COVID, medRxiv, ELSEVIER, bioRxiv, LILACS, Grey literature, Lanzhou University/CNKI, ChemRxiv, WPRIM (Western Pacific), SSRN, ProQuest Central, PREPRINT-SCIELO, PubMed, ArXiv (2287)

Appendix 4. Data extraction categories

Study information

- Study ID
- Study title
- Study source (i.e. journal, report, preprint publication)
- Date of submission
- Date of publication

Study type

- Study type (e.g. experimental, quasi-experimental, observational, modelling)
- Verbal summary of study type (e.g. stochastic discrete event simulation model)
- Comments

Population, setting, and context

- Country in which travel-related control measure is implemented
- Region protected by travel-related control measure
- Region restricted by travel-related control measure
- Short description of population studied (e.g. international travelers including flight passenger and crew members arriving between 27 April and 22 June 2009)
- Comments

Pathogen/disease

- Disease (i.e. SARS, MERS, COVID-19)
- Stage of pandemic in the protected and restricted regions during which the intervention was implemented (i.e. no cases, sporadic cases, clusters of cases, and community transmission)

  - Supporting information for the categorization of the stage of pandemic in the protected and restricted regions (e.g. date of notification of first case, current number of cases, established community spread)

- **Reproduction number: basic reproduction number (R0) and current reproduction number R(t) for the region protected by the intervention**

Intervention

- Broad measure category (i.e. closure of national borders to entry and/or exit, international travel restrictions/bans, entry and/or exit screening at national borders, quarantine or isolation of travelers crossing national borders)
- Verbal summary of specific travel-related measure(s) (e.g. all arriving passengers were provided with a symptom questionnaire)
- Mode of travel restricted by the travel-related control measure(s)
- Date(s) of implementation of the travel-related control measure(s)
- Any reported exceptions to the measure (e.g. certain individuals being excluded or fast-tracked in airport screening because of nationality, occupation, country of origin)
- **Representation of the intervention in the model (e.g. 100% travel reduction)**
- Counterfactual/comparison

- Comments

Primary outcomes (repeated for each primary outcome)

- Outcome category (i.e. cases avoided due to the intervention, number of cases detected due to the intervention, shift in epidemic development due to the intervention)
- Description of outcome

- Length of follow-up
- Estimate related to the impact of the travel-related control measure
- Narrative summary of overall impact of travel-related control measure
- **Additional analysis used to quantify the range of potential effects for this outcome**

- Comments

Secondary outcomes (repeated for each secondary outcome)

- Description of any secondary (non-quantitative) outcomes (i.e. any other infectious disease transmission outcomes, healthcare utilisation outcomes, resource requirements for implementing the intervention, and any adverse effects)
- Narrative results of secondary outcomes
- Comments

Appendix 5. QUADAS-2 domains as applied in the rapid review

Domain		Signalling question	Application in this review
Domain 1: participant selection	A. Risk of bias	1.1 Was a consecutive or random sample of participants enrolled?	Assess how the individuals screened and/or quarantined as part of the study were determined; where all individuals were screened (e.g. as part of a blanket screening/quarantine programme at a port of entry or an evacuation flight) or where a random sample was selected, a risk of bias is not likely.
		1.2 Was a case-control design avoided?	If disease status was used to determine the sample, a risk of bias should be considered.
		1.3 Did the study avoid inappropriate exclusions?	Any exclusions to screening/quarantine programmes should be justified; however even with justification, exclusions could lead to bias, especially where the screening and disease status of those excluded are unknown. Thus, if no exclusion criteria were applied, the risk of bias is low.
		Comments	-
		1. Could the selection of participants have introduced bias?	Consider whether bias may have arisen from 1.1-1.3
	B. Concerns regarding applicability	Describe included participants (prior testing, presentation, intended use of index test and setting)	Consider those individuals screened, and whether they are representative of individuals to be screened as part of screening programmes at international borders likely to be used during the COVID-19 pandemic. For example, screening interventions targeting travellers or commuters at sea ports, airports or land borders under regular travelling conditions are often assumed to have a high external validity; while the individuals evacuated from a high-risk region (often from Wuhan, China, in the early phase of the pandemic) are likely to be different from regular travellers and these studies should be therefore regarded as having a low external validity.
		Is there concern that the included participants do not match the review question?	See above

(Continued)

Domain 2: index test(s)	**A. Risk of bias**	2.1 Can we be sure that those identified in index test (true and false positive screening results) were identified by the index test (e.g. automated fever scanner) rather than any other means (e.g. self-reporting)?	Consider how those screened positive were determined – for entry and/or exit symptom/exposure-based screening all 'positives' should stem from the symptom screening (e.g. a febrile COVID-19 was who was identified by a thermal imaging system at an airport), and not from any other procedures (e.g. self-reporting of cases missed by the screening intervention; based on respiratory symptoms).
		2.2 Were the index test results interpreted without knowledge of the results of the reference standard?	Consider whether, for example, the results of the PCR test were known when symptom or fever screening was applied to individuals.
		2.3 If a threshold was used, was it prespecified?	Consider for temperature screening, whether the cut-off for determining acceptable/high temperature was predefined; for symptom screening, consider whether any symptom or a certain threshold of symptoms was used in defining whether an individual was symptomatic and whether this was predefined.
		Comments on risk of bias	-
		2. Could the conduct or interpretation of the index test have introduced bias?	Consider whether bias may have arisen from 2.1-2.3
	B. Concerns regarding applicability	Describe the index test and how it was conducted and interpreted	Consider the screening/quarantine program assessed, and whether it is representative of one likely to be applied as part of screening programs at international borders during the COVID-19 pandemic. For example, screening interventions such as thermal imaging systems implemented at an airport to detect febrile individuals are assumed to have a high external validity; while the comprehensive medical examinations and observations conducted as part of the repatriation studies (studies in which individuals evacuated from a high risk region (often from Wuhan, China, in the early phase of the pandemic), which were not part of routine border crossing screening, should be regarded as having low external validity
		Is there concern that the index test, its conduct, or interpretation differ from the review question?	See above
Domain 3: approach to identify cases and timing	**A. Risk of bias**	3.1 Is the reference standard (the approach to identify and classify 'cases') likely to correctly classify the target condition (is there	Consider whether the approach to identify cases may have missed relevant cases or classified individuals not infected with SARS-CoV-2 as a case. Any method other than positive PCR test results can be considered at high risk of bias.

		active infection with SARS-CoV-2)?	For the studies using a case-classification based on a positive PCR tests, we assumed the risk of bias due to false positives as low due to the high specificity of the PCR test (in particular if the population is assumed to have a high risk of infection).
			However, there is a considerable risk of false negatives for the PCR test, primarily due to the course of infection (e.g. very low probability of detection in the first days after infection), but also due to inadequate procedures for specimen collection, handling, transportation, or storage (e.g. if only a single test shortly after an infection is applied to a swab sample, the viral load in the individual may not have been high enough for detection, leading to a false negative test).
			We therefore assume a high risk of bias in studies, where asymptomatic individuals do not receive at least two PCR tests and symptomatic individuals did not receive at least two PCR tests after symptom onset.
		3.2. Were the reference standard results interpreted without knowledge of the results of the index test?	Consider whether, for example, the results of the symptom screening were known when the classification was conducted. For PCR tests, where the a risk of subjective judgements to have led to a risk of erroneously classifying a test result as negative or positive is regarded as low, this knowledge of the outcome of the index test is still regarded as leading to a low risk of bias.
		Comments on risk of bias	-
		3. Could the reference standard, its conduct, or its interpretation of the have introduced bias?	Consider whether bias may have arisen from 3.1-3.2
		Describe the reference standard and how it was conducted and interpreted	Consider the procedure for determining who receives the reference standard (the PCR test used to identify cases), and whether it is representative of that likely to be applied as part of screening programmes at international borders during the COVID-19 pandemic.
	B. Concerns regarding applicability	Is there concern that the target condition as defined by the reference standard does not match the review question?	See above
Domain 4: flow and timing		4.1. Did all passengers receive the reference standard?	Consider whether all individuals received the reference test (the respective approach to identify and classify 'cases'; in most cases likely the PCR test).
			For example, if only those who were screened positive (positive index test) and those who developed symptoms during a quarantine observational period were given a PCR test, as this would have led to a high risk of bias due to cases being missed).

(Continued)

			If individuals declined to or for other reasons receive the reference standard (e.g. PCR test), this could lead to cases being missed, which puts the study at a high risk of bias.
			Note: this is independent from 3.1, which evaluates the appropriateness of the approach to classify individuals as cases.
	A. Risk of bias	4.2. Did all passengers receive the same reference standard?	Consider whether the procedure for identifying cases was the same across all individuals or whether it was applied differently without an adequate justification (e.g. individuals with symptoms receiving a different testing procedure).
			Studies, which used different approaches for classifying cases (e.g. some cases defined based on chest computer tomography and some based on PCR) would be classified as high risk of bias.
			Studies in which the classification of cases is based on multiple PCR tests, we consider a high risk of bias if some symptomatic individuals were treated differently from other symptomatic individuals (e.g. some received more PCR tests than others) and if some of the asymptomatic individuals were treated differently from asymptomatic individuals.
		4.3. Were all passengers included in the analysis? Is there likely no or a very low risk of attrition bias?	Consider whether some individuals may have been excluded from the analysis, this would lead to a high risk of bias.
		4.4. Is it possible that the true disease status could have changed between the application of the index test and the reference standard? (e.g. additional infections through how the quarantine was handled?)	Consider whether individuals may have become infected after the initial screening, e.g. if being quarantined among other infected individuals led to some initially non-infected individuals becoming infected. If there is a high risk that individuals who were classified as cases were not cases (i.e. not infected with SARS-CoV-2) at the time when the index test was applied, this would lead to a high risk of bias.
		Comments on risk of bias	-
		4. Could the passenger flow have introduced bias?	Consider whether bias may have arisen from 4.1-4.4

Appendix 6. Criteria used for assessing the quality of individual modelling studies

Aspect	Source	Questions	Application in this review	Examples

(Continued)

Model structure	Philips 2006	1. Are the structural assumptions transparent and justified?	1. Assess whether all structural model assumptions are explicitly stated and whether the authors substantiate these assumptions either through theoretical reasoning or through prior knowledge from the literature.	• Description of model type and defining equations • Comprehensible explanation of model variables and equations • Description of features of the disease captured by the model, e.g. a randomly distributed incubation time • Explanations of model structure implications by text or graphical representations visualising the simulation pathway, e.g. a scheme of the context being modelled • Description of model limitations and simplifying assumptions.
		2. Are the structural assumptions reasonable given the overall objective, perspective and scope of the model?	2. Consider whether the structural assumptions are consistent with what is known about the phenomenon of interest in the literature. In case of disagreement, assess to what extent these discrepancies undermine the overall validity of results and conclusions.	
Input data	Caro 2014	3. Are the input parameters transparent and justified?	3. Assess whether the values of all input parameters are explicitly stated and whether the authors substantiate these values either through theoretical reasoning or through prior knowledge from the literature.	• Epidemiological characteristics known from other studies • Inputs to data calibration algorithms • Table with input parameters and probability distributions used for probabilistic modelling • Explanation and discussion of choice of parameter values with appropriate citations
		4. Are the input parameters reasonable?	4. Consider whether the input parameter values are consistent with what is known about the phenomenon of interest in the literature. In case of disagreement, assess to what extent these discrepancies undermine the overall validity of results and conclusions.	
Validation (external)	Caro 2014	5. Has the external validation process been described?	5. Assess whether there was a formal process of comparing the predictions of the model with 1) the data source that was used to build the model (dependent validation), 2) a data source that was not used to build the model, e.g. an independent country (independent validation) or 3) future values that did not intervene in model building (predictive validation).	• Calibration of SEIR model to case data (dependent validation) • Prediction of a subset of observed data points based on training data set and comparison with validation data set (dependent validation) • Prediction of data points of country/region that was not part of the model fitting and calibration process and comparison with observed data (independent validation) • Prediction of future values that were not used in model building (predictive validation)
		6. Has the model been shown to be externally valid?	6. Consider the extent to which model predictions agree with the data sources that were selected for the external validation process.	

(Continued)

Validation (internal)	Caro 2014	7. Has the internal validation process been described?	7. Assess whether there was a formal process of verifying the extent to which the mathematical calculations are consistent with the model's specifications, e.g. in the form of a simulation study in which the mathematical calculations are applied to data that were simulated according to the model with known parameter values.	• Application of the model on simulated data to establish that analyses work as intended • Code review process conducted by authors or by an independent source to ensure correct implementation of mathematical structure • Independent replication of model
		8. Has the model been shown to be internally valid?	8. Consider the extent to which the results of the internal validation process indicate that the mathematical calculations are consistent with the model's specifications.	
Uncertainty	Caro 2014	9. Was there an adequate assessment of the effects of uncertainty?	9. Consider whether the robustness of results to alternative input parameter values or model assumptions was assessed either by reporting the results of specific sensitivity analyses or through an app in which readers can themselves explore the effects of varying these model assumptions and input parameter values.	• Structural and parameter sensitivity analyses • Inherent stochasticity due to simulation nature of model • Reporting of an app in which effects of input changes can be tracked • Propagation of present uncertainties to outcomes • Was the model probabilistic, i.e. were parameter values fixed or sampled from a distribution? • Is uncertainty transparently reported, described and justified?
Transparency	Caro 2014	10. Was technical documentation, in sufficient detail to allow (potentially) for replication, made available openly or under agreements that protect intellectual property?	10. Assess whether the description of the analyses (including model structure, input parameters, data sources and methods) is sufficiently detailed to allow for the replication of results. In particular, consider whether the code that was used to obtain the results is freely available and well documented.	• Description of model which is qualitatively extensive enough to allow for scrutiny of other researchers (e.g. supplementary material) • Do authors encourage replication by clarifying a procedure to obtain code? • Do the authors only refer to other, similar models for justification and detailed methodological description or do they provide their own documentation?

Appendix 7. Studies excluded from the analysis

Twenty-two studies were identified which met the inclusion criteria and assessed travel-related control measures, but which did not contain sufficient data to inform decision-making.

We identified eight observational studies assessing entry and/or exit screening measures and reporting only limited data on the effectiveness of the measures (Chang 2020; Expert-Taskforce 2020; Gupta 2020; Hayakawa 2020; Ing 2020; Jernigan 2020; Potdar 2020; Sriwijitalai 2020a). Three studies described screening and management of travellers on cruise ships with data on the number of

cases detected, number of individuals screening positive and proportion of deaths among the passengers and the crew (Expert-Taskforce 2020; Gupta 2020; Ing 2020). Another study also presented data on the number of cases detected due to entry symptom screening and testing measures in a cohort of individuals repatriated by air to Japan from China (Hayakawa 2020). Four studies examined entry screening and quarantine measures for travellers crossing national borders. These studies reported data on travellers entering India (Potdar 2020), Taiwan (Chang 2020), Thailand (Sriwijitalai 2020a), and the USA (Jernigan 2020); the studies reported on the number or proportion of cases detected, but did not provide data on how many cases were missed. In general, these studies showed that entry and/or exit screening measures and quarantine of travellers allows for the detection of cases in a variety of settings, however the effectiveness of the measures could not be further assessed with the reported data.

Nine observational ecological studies examined the effects of travel restrictions reducing or stopping cross-border travel (Arshed 2020; Chaudhry 2020; Jablonska 2020; Koh 2020; Leffler 2020; Liu 2020a; Ogundokun 2020; Stokes 2020; Teixeira da Silva 2020). These studies provided aggregated data on the impact of the measures in this intervention category across various countries. Most of these studies reported aggregated data on more than 100 countries implementing the measures (Arshed 2020; Koh 2020; Leffler 2020; Liu 2020a; Stokes 2020; Teixeira da Silva 2020), and only one focused on Nigeria, specifically (Ogundokun 2020). The regions restricted by the measures also varied and were often not reported specifically. The studies reported various outcomes related to the cases avoided due to the measure (Chaudhry 2020; Leffler 2020; Ogundokun 2020; Stokes 2020; Teixeira da Silva 2020), and the shift in epidemic development (Arshed 2020; Jablonska 2020; Koh 2020; Liu 2020a). In general, the results from these studies reflect those of the main studies in this intervention category showing beneficial effects on the cases avoided and in shifting the epidemic development, with some suggestion that earlier implementation may be more effective.

Five modelling studies examined the effects of travel restrictions reducing or stopping cross-border travel (Baba 2020; Cacciapaglia 2020a; Cacciapaglia 2020b; Chen 2020d; Jorritsma 2020). Four studies reported hypothetical regions for the measure implementation and restriction (Baba 2020; Cacciapaglia 2020b; Chen 2020d; Jorritsma 2020), and one study simulated the measures in European countries (Cacciapaglia 2020a). One of the studies reported that reductions in travel into a region leads to a reduction in cumulative cases (Chen 2020d), another study reported on cases avoided due to the measure and found no beneficial effect (Baba 2020), and three studies reported on the shift in epidemic development predicting delays in the outbreak, specifically when measures were implemented early (Cacciapaglia 2020a; Cacciapaglia 2020b; Jorritsma 2020).

Characteristics of supporting studies

Study ID	Country implementing the measure	Country restricted by the measure	Intervention category	Date of implementation	Outcome
Observational screening studies					
Chang 2020	Taiwan	COVID-19 epidemic areas	Entry screening at the airports	26 February to 17 March 2020	Proportion of cases detected
Expert-Taskforce 2020	Cruise ship quarantined in Japan	n.r.	Symptom screening followed by testing and disembarking of those being positive	4 February 2020	Number of cases detected; number of asymptomatic cases detected; proportion of deaths
Gupta 2020	Cruise ship quarantined in Japan	n.r.	Symptom screening followed by testing and disembarking of those being positive	4 February 2020	Number of cases detected; number of symptomatic cases detected; proportion of deaths; number of cases requiring intensive care
Jernigan 2020	USA	China	Entry symptom screening; testing; 14-day quarantine of those being negative	n.r.	Number of individuals screening positive; number of cases detected
Hayakawa 2020	Japan	China	Entry symptom screening; testing	29 January 2020 to 17 February 2020	Number of individuals screening positive; number of cases detected

Cochrane
Library

Trusted evidence.
Informed decisions.
Better health.

Cochrane Database of Systematic Reviews

(Continued)

Ing 2020	Cruise ship departed from Argentina	n.r.	Symptom screening; testing; isolation; medical evacuation	Day 8 of the cruise	Number of individuals screening positive; number of cases; number of deaths
Potdar 2020	India	All countries	Entry testing	22 January to 29 February 2020	Number of cases detected
Sriwijitalai 2020a	Thailand	Unclear whether screening applies to all arrivals, or only those from China	Entry symptom screening	Since early December 2019	Number of individuals screened positive
Observational ecological studies					
Arshed 2020	190 countries	n.r.	Travel restrictions reducing or stopping cross-border travel	22 January to 11 May 2020	Time to outbreak
Chaudhry 2020	50 countries with the highest number of cases as of May 1, 2020	n.r.	Travel restrictions reducing or stopping cross-border travel	By 1 April 2020	Number of cases in the community; number of deaths; number of cases recovered; number of critical cases (all outcomes measured as 'days to the restrictions')
Jablonska 2020	34 European countries	n.r.	Travel restrictions reducing or stopping cross-border travel	The date of border closures in each country	Time to outbreak; number of deaths at peak
Koh 2020	142 countries	n.r.	Travel restrictions reducing or stopping cross-border travel	n.r.	Effective reproduction number
Leffler 2020	200 countries	n.r.	Travel restrictions reducing or stopping cross-border travel	By 16 April 2020	Number of deaths (per capita)
Liu 2020a	130 countries	n.r.	Travel restrictions reducing or stopping cross-border travel	By 22 June 2020	Effective reproduction number
Ogundokun 2020	Nigeria	All international travellers	Travel restrictions reducing or stopping cross-border travel	14 April 2020	Number of cases in the community
Teixeira da Silva 2020	121 countries	n.r.	Travel restrictions reducing or stopping cross-border travel	n.r.	Number of deaths (per cases)
Stokes 2020	130 countries	n.r.	Travel restrictions reducing or stopping cross-border travel	n.r.	Number of deaths
Modelling studies					

(Continued)

Baba 2020	Hypothetical	Hypothetical	Travel restrictions reducing or stopping cross-border travel	n.r.	Number of cases in the community
Cacciapaglia 2020a	European countries	n.r.	Travel restrictions reducing or stopping cross-border travel	Starting in week 25 of 2020	Time to outbreak
Cacciapaglia 2020b	Hypothetical	Hypothetical	Travel restrictions reducing or stopping cross-border travel	n.r.	Time to outbreak
Chen 2020d	Hypothetical	Hypothetical	Travel restrictions reducing or stopping cross-border travel	n.r.	Number of cumulative cases
Jorritsma 2020	Hypothetical	Hypothetical	Travel restrictions reducing or stopping cross-border travel	n.r.	Number of cases at peak

Appendix 8. QUADAS-2 Risk of bias and applicability assessment of observational screening studies

Study ID	Domain	Signalling question	Rating
Symptom screening			
Al-Qahtani 2020	Domain 1: Patient selection - A. Risk of bias	1.1 Was a consecutive or random sample of participants enrolled?	**Yes**
		1.2 Was a case-control design avoided?	**Yes**
		1.3 Did the study avoid inappropriate exclusions?	**Yes**
		Comments	2714 individuals arriving to Bahrain airport underwent PCR testing and quarantine procedures
		1. Could the selection of participants have introduced bias?	**Low**
	Domain 1: Patient selection - B. Concerns regarding applicability	Describe included participants (prior testing, presentation, intended use of index test and setting):	The study reports on individuals travelling to Bahrain between 25 February and 14 March 2020. As it is not a repatriation study, the population is close to real world travellers. However, the transferability of the travelling population in the early phase of the pandemic (those returning at times of border closure initiation in numerous countries) could be regarded as different than regular travel
		Is there concern that the included participants do not match the review question?	**Low**
	Domain 2: Index test(s)	2.1 Can we be sure that those identified in index test (true and false pos-	**Unclear**

Cochrane
Library

Trusted evidence.
Informed decisions.
Better health.

Cochrane Database of Systematic Reviews

(Continued)

	A. Risk of bias	itive screening results) were identified by the index test (e.g. automated fever scanner) rather than any other means (e.g. self-reporting)?	
		2.2 Were the index test results interpreted without knowledge of the results of the reference standard?	**Unclear**
		2.3 If a threshold was used, was it prespecified?	**Unclear**
		Comments on risk of bias	The measure of only testing those individuals with a positive PCR test is a hypothetical intervention. The individuals received a symptom based screening and where classified as symptomatic or non-symptomatic. It is likely, that the status of "symptomatic" was defined prior to conducting the PCR test, but this is not clear.
		2. Could the conduct or interpretation of the index test have introduced bias?	**Unclear**
	Domain 2: Index test(s) B. Concerns regarding applicability	Describe the index test and how it was conducted and interpreted:	The symptom/exposure-based screening was conducted on regular travellers travelling to Bahrain; It can be therefore assumed to be - while resource intensive - a regular intervention. However the travel volume is lower than would likely be expected under real world scenarios in most situations without additional travel restrictions / border closures
		Is there concern that the index test, its conduct, or interpretation differ from the review question?	**Unclear**
	Domain 3: Reference standard A. Risk of bias	3.1 Is the reference standard (the approach to identify and classify 'cases') likely to correctly classify the target condition (here active infection with SARS-CoV-2)?	**Yes**
		3.2. Were the reference standard results interpreted without knowledge of the results of the index test?	**No**
		Comments on risk of bias	As those with a negative initial PCR test underwent quarantine and retesting prior to release from quarantine, the status of the individuals was known to the testers. But as all individuals received the test and PCR testing can be considered relatively objective, this is regarded as low risk of bias. The risk of status change is discussed in domain 4.

	3. Could the reference standard, its conduct, or its interpretation have introduced bias? (Leading to an under-determination or over-determination of true findings in the reference test)	**Low**
Domain 3: Reference standard & timing B. Concerns regarding applicability	Describe the reference standard and how it was conducted and interpreted:	Reference test was Observation during 14 day quarantine-period with retesting in the case of symptoms. and second PCR test for all asymptomatic passengers with an initial negative first PCR test and no symptoms during quarantine. While there is a risk of two negative PCR tests in a row without the development of symptoms, this risk is low.
	Is there concern that the target condition as defined by the reference standard does not match the review question?	**Low**
Domain 4: Flow and timing - Risk of bias	4.1 Did all passengers receive the reference standard?	**Yes**
	4.2 Did all passengers receive the same reference standard?	**Yes**
	4.3 Were all passengers included in the analysis? Is there likely no or a very low risk of attrition bias?	**Yes**
	4.4. Is it possible that the true disease status could have changed between the application of the index test and the reference standard? (e.g. additional infections through how the quarantine was handled?)	**Unclear**
	Comments on risk of bias	All Individuals received the PCR test at the same point of time for each repatriation flight and by the same group of experts. There could be a low risk of between flights contamination; e.g. due to differences in taking the swab between experts. The risk of infection during the quarantine period is unclear, this is not well described. There is a risk of individuals having a first false negative PCR test. During the 14-day quarantine period, those individuals cure out the infection. At the second PCR test they are correctly tested negative. There is a low risk of the time delay between the index test and the reference test leads to missing infected individuals within the population of those with two negative PCR tests.

(Continued)

		4. Could the passenger flow have introduced bias?	**Unclear**
Arima 2020	Domain 1: Patient selection - A. Risk of bias	1.1 Was a consecutive or random sample of participants enrolled?	**Yes**
		1.2 Was a case-control design avoided?	**Yes**
		1.3 Did the study avoid inappropriate exclusions?	**Unclear**
		Comments	All 566 passengers repatriated were screened and followed-up over the 14-day quarantine period. 3 symptomatic passengers were denied boarding. Their status is unclear. This could lead to an underestimation of the effectiveness of symptom/exposure-based screening
		1. Could the selection of participants have introduced bias?	**Unclear**
	Domain 1: Patient selection - B. Concerns regarding applicability	Describe included participants (prior testing, presentation, intended use of index test and setting):	The population is a highly selected sample in the early phase of the pandemic and does not represent regular air travel passengers which would travel between China and Japan
		Is there concern that the included participants do not match the review question?	**High**
	Domain 2: Index test(s) A. Risk of bias	2.1 Can we be sure that those identified in index test (true and false positive screening results) were identified by the index test (e.g. automated fever scanner) rather than any other means (e.g. self-reporting)?	**Yes**
		2.2 Were the index test results interpreted without knowledge of the results of the reference standard?	**Yes**
		2.3 If a threshold was used, was it prespecified?	**Unclear**
		Comments on risk of bias	
		2. Could the conduct or interpretation of the index test have introduced bias?	**Low**
	Domain 2: Index test(s) B. Concerns regarding applicability	Describe the index test and how it was conducted and interpreted:	Repatriation study, real-world generalisability should be considered. Likely, the awareness of the risk status of the individuals led to led to higher vigilance than would take place in a normal airport screening.

(Continued)

			"Day 1 entry screening by testing oropharyngeal swab samples collected from all 566 returnees at the hospitals to which they were initially transported for SARS-CoV-2 (4); all tests were based on the real-time reverse transcription PCR developed by the National Institute of Infectious Diseases"
		Is there concern that the index test, its conduct, or interpretation differ from the review question?	**High**
	Domain 3: Reference standard A. Risk of bias	3.1 Is the reference standard (the approach to identify and classify 'cases') likely to correctly classify the target condition (here active infection with SARS-CoV-2)?	**Yes**
		3.2. Were the reference standard results interpreted without knowledge of the results of the index test?	**Unclear**
		Comments on risk of bias	Reference test was given at both entry and 14 days later at quarantine exit; there is a slight chance given the incubation and infectiousness periods that some cases were not captured, however this risk of rather low.
		3. Could the reference standard, its conduct, or its interpretation have introduced bias? (Leading to an under-determination or over-determination of true findings in the reference test)	**Low**
	Domain 3: Reference standard & timing B. Concerns regarding applicability	Describe the reference standard and how it was conducted and interpreted:	Exit screening for quarantined persons who remained illness-free (i.e. not already identified as a case) by collecting oropharyngeal swab samples on day 14, the end of the quarantine period.
		Is there concern that the target condition as defined by the reference standard does not match the review question?	**Low**
	Domain 4: Flow and timing - Risk of bias	4.1 Did all passengers receive the reference standard?	**Yes**
		4.2 Did all passengers receive the same reference standard?	**Yes**
		4.3 Were all passengers included in the analysis? Is there likely no or a very low risk of attrition bias?	**Yes**

(Continued)

		4.4. Is it possible that the true disease status could have changed between the application of the index test and the reference standard? (e.g. additional infections through how the quarantine was handled?)	**Unclear**
		Comments on risk of bias	All passengers received the reference test at least once at entry and once at quarantine exit
		4. Could the passenger flow have introduced bias?	**Unclear**
Chen J 2020	Domain 1: Patient selection - A. Risk of bias	1.1 Was a consecutive or random sample of participants enrolled?	**Yes**
		1.2 Was a case-control design avoided?	**Yes**
		1.3 Did the study avoid inappropriate exclusions?	**Yes**
		Comments	All passengers on the flight were included in the study. The crew was not included in the analysis.
		1. Could the selection of participants have introduced bias?	**Low**
	Domain 1: Patient selection - B. Concerns regarding applicability	Describe included participants (prior testing, presentation, intended use of index test and setting):	The population is a highly selected sample of travelers where an assumed outbreak on the airplane had occurred
		Is there concern that the included participants do not match the review question?	**High**
	Domain 2: Index test(s) A. Risk of bias	2.1 Can we be sure that those identified in index test (true and false positive screening results) were identified by the index test (e.g. automated fever scanner) rather than any other means (e.g. self-reporting)?	**Yes**
		2.2 Were the index test results interpreted without knowledge of the results of the reference standard?	**Yes**
		2.3 If a threshold was used, was it prespecified?	**Unclear**
		Comments on risk of bias	Symptoms of relevance and definition for classification as "symptomatic" are provided. Screening was conducted at the airport quarantine inspection. This is a hypothetical intervention; all individuals were tested with a PCR upon arrival;

(Continued)

			and those being symptomatic were reported. It cannot be ruled out that other approaches would have been conducted; if only those with symptoms were tested
	2. Could the conduct or interpretation of the index test have introduced bias?	**Unclear**	
Domain 2: Index test(s) B. Concerns regarding applicability	Describe the index test and how it was conducted and interpreted:	The screening procedure was conducted based on several passengers returning from Wuhan and 2 people being found to be symptomatic. Likely, this led to higher vigilance than would take place in a normal airport screening	
	Is there concern that the index test, its conduct, or interpretation differ from the review question?	**High**	
Domain 3: Reference standard A. Risk of bias	3.1 Is the reference standard (the approach to identify and classify 'cases') likely to correctly classify the target condition (here active infection with SARS-CoV-2)?	**yes**	
	3.2. Were the reference standard results interpreted without knowledge of the results of the index test?	**No**	
	Comments on risk of bias	All participants received two PCR tests + symptom observation with PCR testing as reference test. it is likely that the result of this testing regime captures most infected individuals. Participants were diagnosed using PCR test kits recommended by the Chinese CDC in accordance with protocols established by the WHO. The symptom observation during quarantine was conducted by temperature measurement twice daily and report the development of symptoms.	
	3. Could the reference standard, its conduct, or its interpretation have introduced bias? (Leading to an under-determination or over-determination of true findings in the reference test)	**Low**	
Domain 3: Reference standard & timing B. Concerns regarding applicability	Describe the reference standard and how it was conducted and interpreted:	All participants received two PCR tests at day 0 and day 13 and a symptom observation with temperature measurement twice daily with PCR testing upon symptom development was used as reference test.	
	Is there concern that the target condition as defined by the reference	**Low**	

(Continued)

		standard does not match the review question?	
	Domain 4: Flow and timing - Risk of bias	4.1 Did all passengers receive the reference standard?	**Yes**
		4.2 Did all passengers receive the same reference standard?	**Yes**
		4.3 Were all passengers included in the analysis? Is there likely no or a very low risk of attrition bias?	**Yes**
		4.4. Is it possible that the true disease status could have changed between the application of the index test and the reference standard? (e.g. additional infections through how the quarantine was handled?)	**Unclear**
		Comments on risk of bias	Other than children with families, travellers were quarantined in single rooms in the hotel. There are limited descriptions on the quarantine regulation (e.g. on whether they were allowed to move freely). 3 individuals were identified at day 7 and day 9 after arrival. Two infected individuals were a couple with departure in Wuhan and belonging to the same tour group as most of the infected, including the likely two index participants. One individual with a symptom development at day 9 after arrival had a departure in Wuhan and did not belong to the tour group as most of the infected. It cannot be ruled out that those tested positive had developed symptoms during the quarantine process without additional information.
		4. Could the passenger flow have introduced bias?	**Unclear**
Hoehl 2020	Domain 1: Patient selection - A. Risk of bias	1.1 Was a consecutive or random sample of participants enrolled?	**Yes**
		1.2 Was a case-control design avoided?	**Yes**
		1.3 Did the study avoid inappropriate exclusions?	**Yes**
		Comments	All 126 passengers on board the aircraft from Wuhan received the index text. Likely low risk of selection bias
		1. Could the selection of participants have introduced bias?	**Low**

(Continued)

	Domain 1: Patient selection - B. Concerns regarding applicability	Describe included participants (prior testing, presentation, intended use of index test and setting):	The population is a highly selected sample in the early phase of the pandemic and does not represent regular air travel passengers which would travel between China and Germany
		Is there concern that the included participants do not match the review question?	**High**
	Domain 2: Index test(s) A. Risk of bias	2.1 Can we be sure that those identified in index test (true and false positive screening results) were identified by the index test (e.g. automated fever scanner) rather than any other means (e.g. self-reporting)?	**Yes**
		2.2 Were the index test results interpreted without knowledge of the results of the reference standard?	**Yes**
		2.3 If a threshold was used, was it prespecified?	**Unclear**
		Comments on risk of bias	The study does not provide information on the first screening in China and does not explicitly state that the measures were predefined. But this seems to be likely.
		2. Could the conduct or interpretation of the index test have introduced bias?	**Low**
	Domain 2: Index test(s) B. Concerns regarding applicability	Describe the index test and how it was conducted and interpreted:	The two step screening procedure was not a regular airport screening, but rather an intervention which is similar to a symptom- and exposure based screening that could be conducted in the same way. Likely, the awareness of the risk status of the individuals led to led to higher vigilance than would take place in a normal airport screening. "Screening for symptoms and clinical signs of infection was performed before their departure from China" "During the flight, 10 passengers were isolated. "These 10 passengers were transferred to University Hospital Frankfurt immediately after arrival." "Two passengers had had contact with 1 person who had a confirmed case of SARS-CoV-2 infection, 6 had reported symptoms, were deemed to be clinically symptomatic, or both, and 2 passengers had accompanied family members who had been isolated on the flight because of suspected SARS-CoV-2 infection or because of other symptoms (i.e., symptoms related to pregnancy)." "The remaining 116 passengers [...] were sent to the medical assessment center at Frankfurt Airport, where each was evaluated by a med-

![Cochrane Library logo] Trusted evidence.
Informed decisions.
Better health.

Cochrane Database of Systematic Reviews

(Continued)

		ical team of physicians. Each passenger was asked to report current symptoms of fever, fatigue, sore throat, cough, runny nose, muscle aches, and diarrhea, and each one was screened for signs of infection in the nose and throat. The temperature of all passengers was taken."
	Is there concern that the index test, its conduct, or interpretation differ from the review question?	**High**
Domain 3: Reference standard A. Risk of bias	3.1 Is the reference standard (the approach to identify and classify 'cases') likely to correctly classify the target condition (here active infection with SARS-CoV-2)?	**No**
	3.2. Were the reference standard results interpreted without knowledge of the results of the index test?	**Unclear**
	Comments on risk of bias	1. Likelihood of having missed a positive finding in reference test: Unclear. - Participants were isolated for 14 days; there should have been a chance of around 70% to show symptoms if a passenger was infected. - All but 1 passenger received an RT-PCR test. However, limited information is provided to judge if there is a risk of bias due to (a) inadequate taking of the sample (likely low, if conducted by medical professional) and (b) inadequate transportation, storage, or delay in testing (likely low, but unclear). - It is unclear, (a) at what point of time during the quarantine the test was conducted, (b) if it was repeated (2 tests per person), and (c) if the repetition of the two tests (if there were 2) had a sufficient delay. 2. likelihood of having misclassified a positive case in reference test: low, cases received three tests (at least 2 PCR tests and 1 cell culture test)
	3. Could the reference standard, its conduct, or its interpretation have introduced bias? (Leading to an under-determination or over-determination of true findings in the reference test)	**Unclear**
Domain 3: Reference standard & timing B. Concerns regarding applicability	Describe the reference standard and how it was conducted and interpreted:	All (n = 11) passengers with positive screening results received a PCR; For those (n = 116) remaining passengers in quarantine, 115 received the test (1 person declined; no reasons provided); Two passengers had a positive test result from RT-PCR which was later confirmed by a second test.

(Continued)

			"RT-PCR (cycle threshold value in the two samples, 24.39 and 30.25, respectively). Testing with a second protocol consisting of two commercial sets (LightMix Modular SARS and Wuhan CoV E-gene, and LightMix Modular Wuhan CoV RdRP-gene, both produced by TIB MOLBIOL) and retesting of the positive samples at the Institute of Virology, Philipps University Marburg, in Marburg, Germany, confirmed the results"
		Is there concern that the target condition as defined by the reference standard does not match the review question?	**Low**
	Domain 4: Flow and timing - Risk of bias	4.1 Did all passengers receive the reference standard?	**Unclear**
		4.2 Did all passengers receive the same reference standard?	**Yes**
		4.3 Were all passengers included in the analysis? Is there likely no or a very low risk of attrition bias?	**Unclear**
		4.4. Is it possible that the true disease status could have changed between the application of the index test and the reference standard? (e.g. additional infections through how the quarantine was handled?)	**Unclear**
		Comments on risk of bias	All but one passenger received the standard test (1/126). The risk of bias resulting from this is considered low, but the reasons for rejecting the reference test are unclear. Limited information on the quarantine procedure is reported
		4. Could the passenger flow have introduced bias?	**Unclear**
Kim 2020	Domain 1: Patient selection - A. Risk of bias	1.1 Was a consecutive or random sample of participants enrolled?	**Yes**
		1.2 Was a case-control design avoided?	**Yes**
		1.3 Did the study avoid inappropriate exclusions?	**Yes**
		Comments	All passengers on the flight were included; there is no indication that there was a preselection of individuals

(Continued)

	1. Could the selection of participants have introduced bias?	**Low**
Domain 1: Patient selection - B. Concerns regarding applicability	Describe included participants (prior testing, presentation, intended use of index test and setting):	The population is a highly selected sample of repatriated individuals in the early phase of the pandemic and does not represent regular air travel passengers,
	Is there concern that the included participants do not match the review question?	**High**
Domain 2: Index test(s) A. Risk of bias	2.1 Can we be sure that those identified in index test (true and false positive screening results) were identified by the index test (e.g. automated fever scanner) rather than any other means (e.g. self-reporting)?	**Yes**
	2.2 Were the index test results interpreted without knowledge of the results of the reference standard?	**Yes**
	2.3 If a threshold was used, was it prespecified?	**Yes**
	Comments on risk of bias	Individuals were identified based on a fixed procedure regarding symptom identification; if action based on the existing criteria (e.g. hospitalisation of those who were screened positive)
	2. Could the conduct or interpretation of the index test have introduced bias?	**Low**
Domain 2: Index test(s) B. Concerns regarding applicability	Describe the index test and how it was conducted and interpreted:	As it was a relatively small sample, it is unclear if symptom screening with the same intensity could be conducted under real world circumstances of regular air travel. "...the pre-filled Health Status Questionnaires (HSQs) and passengers' body temperature were taken. Passengers who reported respiratory symptoms in the previous 14 days on the HSQ or who had a temperature of ≥ 37.5°C were immediately assessed for COVID-19 by the quarantine doctor. Passengers had a blue or a red sticker placed on their chest, depending on whether they were asymptomatic or under investigation, respectively, so that their COVID-19 status could be easily identified. [...] [Arrival screening:] The waiting quarantine officers checked the pre-filled HSQs, measured passengers' body temperatures, and asked them about symptoms they experienced during the flight. Based on this screening, 1 addi-

(Continued)

			tional passenger was categorized as a PUI on arrival."
		Is there concern that the index test, its conduct, or interpretation differ from the review question?	**Unclear**
	Domain 3: Reference standard A. Risk of bias	3.1 Is the reference standard (the approach to identify and classify 'cases') likely to correctly classify the target condition (here active infection with SARS-CoV-2)?	**yes**
		3.2. Were the reference standard results interpreted without knowledge of the results of the index test?	**No**
		Comments on risk of bias	The reference test consisted of 1 PCR test upon arrival and symptom observation. PCR tests have a risk of producing false negative results, in particular in the early days of infection. There is a risk of the testing procedure missing individuals who got infected around the days of repatriation and did not develop symptoms, and/or had long incubation periods
		3. Could the reference standard, its conduct, or its interpretation have introduced bias? (Leading to an under-determination or over-determination of true findings in the reference test)	**Unclear**
	Domain 3: Reference standard & timing B. Concerns regarding applicability	Describe the reference standard and how it was conducted and interpreted:	PCR tests are considered the gold standard for identifying SARS-CoV-2 Infections
		Is there concern that the target condition as defined by the reference standard does not match the review question?	**Low**
	Domain 4: Flow and timing - Risk of bias	4.1 Did all passengers receive the reference standard?	**Yes**
		4.2 Did all passengers receive the same reference standard?	**Yes**
		4.3 Were all passengers included in the analysis? Is there likely no or a very low risk of attrition bias?	**Yes**

Cochrane
Library

Trusted evidence.
Informed decisions.
Better health.

Cochrane Database of Systematic Reviews

(Continued)

		4.4. Is it possible that the true disease status could have changed between the application of the index test and the reference standard? (e.g. additional infections through how the quarantine was handled?)	**Unclear**
		Comments on risk of bias	The quarantine procedure is not described in detail; the risk of infections during the quarantine would be classified as unclear. But as there were no additional cases identified afterwards; it is regarded as low.
		4. Could the passenger flow have introduced bias?	**Low**
Lytras 2020	Domain 1: Patient selection - A. Risk of bias	1.1 Was a consecutive or random sample of participants enrolled?	**Unclear**
		1.2 Was a case-control design avoided?	**Yes**
		1.3 Did the study avoid inappropriate exclusions?	**Yes**
		Comments	All passengers on board were screened; Out of all repatriation flights, it is unclear why these 7 selected flights were chosen.
		1. Could the selection of participants have introduced bias?	**Low**
	Domain 1: Patient selection - B. Concerns regarding applicability	Describe included participants (prior testing, presentation, intended use of index test and setting):	The population is a highly selected sample in the early phase of the pandemic and does not represent regular air travel passengers which would travel between Spain, UK, Turkey and Greece
		Is there concern that the included participants do not match the review question?	**High**
	Domain 2: Index test(s) A. Risk of bias	2.1 Can we be sure that those identified in index test (true and false positive screening results) were identified by the index test (e.g. automated fever scanner) rather than any other means (e.g. self-reporting)?	**Unclear**
		2.2 Were the index test results interpreted without knowledge of the results of the reference standard?	**Yes**
		2.3 If a threshold was used, was it prespecified?	**Unclear**
		Comments on risk of bias	Participants reported symptoms and received some form of medical examination. It is un-

(Continued)

			clear how the screening was handled. A swab was taken at the same time, therefore prior to knowing about the results of the reference test
		2. Could the conduct or interpretation of the index test have introduced bias?	**Unclear**
	Domain 2: Index test(s) B. Concerns regarding applicability	Describe the index test and how it was conducted and interpreted:	The symptom/exposure-based screening and how individual cases were handled is not well described. We do not know much about the screening procedure (index test). Likely, the awareness of the risk status of the individuals led to led to higher vigilance than would take place in a normal airport screening; "All passengers consented to screening, and were asked in-flight to fill in a paper form with demographic, clinical and contact information. A temporary facility was set up by NPHO at a gate in Athens airport 'Eleftherios Venizelos', and swab samples were obtained from passengers immediately upon arrival; those not in need of medical care were subsequently requested to self-quarantine at home for 14 days."
		Is there concern that the index test, its conduct, or interpretation differ from the review question?	**High**
	Domain 3: Reference standard A. Risk of bias	3.1 Is the reference standard (the approach to identify and classify 'cases') likely to correctly classify the target condition (here active infection with SARS-CoV-2)?	**No**
		3.2. Were the reference standard results interpreted without knowledge of the results of the index test?	**Unclear**
		Comments on risk of bias	One-Time PCR test upon arrival. No additional reference tests are reported, even among those 36 who developed symptoms in the course of the quarantine. Limited information on how the swab was taken is reported.
		3. Could the reference standard, its conduct, or its interpretation have introduced bias? (Leading to an under-determination or over-determination of true findings in the reference test)	**High**
	Domain 3: Reference standard & timing B. Concerns regarding applicability	Describe the reference standard and how it was conducted and interpreted:	All passengers received the screening test and reference test upon arrival. Only one reference test PCR was conducted, leading to the risk of false negatives. An additional 36 passengers in quarantine with negative tests had

(Continued)

			developed symptoms in the quarantine. It is possible that a proportion of them was suffering from COVID-19.
		Is there concern that the target condition as defined by the reference standard does not match the review question?	**Low**
	Domain 4: Flow and timing - Risk of bias	4.1 Did all passengers receive the reference standard?	**Yes**
		4.2 Did all passengers receive the same reference standard?	**Yes**
		4.3 Were all passengers included in the analysis? Is there likely no or a very low risk of attrition bias?	**Yes**
		4.4. Is it possible that the true disease status could have changed between the application of the index test and the reference standard? (e.g. additional infections through how the quarantine was handled?)	**Unclear**
		Comments on risk of bias	All passengers received the reference test, there is no indication of a different treatment or of attrition bias
		4. Could the passenger flow have introduced bias?	**Unclear**
Ng 2020	Domain 1: Patient selection - A. Risk of bias	1.1 Was a consecutive or random sample of participants enrolled?	**Yes**
		1.2 Was a case-control design avoided?	**Yes**
		1.3 Did the study avoid inappropriate exclusions?	**no**
		Comments	Here, 3 febrile passengers were removed from the pool and we do not have information on their diagnosis; thus removing infected passengers at this point would lead to an underestimate of the screening
		1. Could the selection of participants have introduced bias?	**High**
	Domain 1: Patient selection - B. Concerns regarding applicability	Describe included participants (prior testing, presentation, intended use of index test and setting):	Repatriation-study, transferability to the real world as a screening intervention is unclear.

(Continued)

		"We followed up on 94 persons who boarded an evacuation flight from Wuhan to Singapore on January 30, 2020."
	Is there concern that the included participants do not match the review question?	**High**
Domain 2: Index test(s) A. Risk of bias	2.1 Can we be sure that those identified in index test (true and false positive screening results) were identified by the index test (e.g. automated fever scanner) rather than any other means (e.g. self-reporting)?	**Yes**
	2.2 Were the index test results interpreted without knowledge of the results of the reference standard?	**Yes**
	2.3 If a threshold was used, was it prespecified?	**Yes**
	Comments on risk of bias	On arrival in Singapore, the passengers underwent repeat screening for body temperature (fever was defined as a body temperature ≥38°C), and 2 persons had a fever. Co-Intervention: 1. Airport Departure fever screening: Screening for body temperature was conducted at check-in and before boarding, and 3 febrile persons were prevented from boarding (no additional information regarding the status of these 3 febrile persons was available) 2. Masks Surgical masks were provided to passengers on board the plane.
	2. Could the conduct or interpretation of the index test have introduced bias?	**Low**
Domain 2: Index test(s) B. Concerns regarding applicability	Describe the index test and how it was conducted and interpreted:	Repatriation-study, transferability to the real world as a screening intervention is unclear. Likely, the awareness of the risk status of the individuals led to led to higher vigilance than would take place in a normal airport screening "The 2 febrile women identified in arrival screening were transferred immediately to a hospital, and they tested positive for SARS-CoV-2 (their clinical course is described in the Supplementary Appendix)"
	Is there concern that the index test, its conduct, or interpretation differ from the review question?	**High**

(Continued)

Domain 3: Reference standard A. Risk of bias	3.1 Is the reference standard (the approach to identify and classify 'cases') likely to correctly classify the target condition (here active infection with SARS-CoV-2)?	**No**
	3.2. Were the reference standard results interpreted without knowledge of the results of the index test?	**no**
	Comments on risk of bias	While the reference test is not perfect, most passengers received at least 2 PCR tests with 3 days in between, were observed intensely for 14 days, and tested symptoms developed.. There is a risk of participants being missed by this procedure, but likely this is low.
	3. Could the reference standard, its conduct, or its interpretation have introduced bias? (Leading to an under-determination or over-determination of true findings in the reference test)	**Low**
Domain 3: Reference standard & timing B. Concerns regarding applicability	Describe the reference standard and how it was conducted and interpreted:	1. "Quarantine and observation for 14 days and checking of symptoms and fever three times daily 2. PCR assessment of those showing symptoms; those with symptoms received only 1 test 3. On quarantine day 3, samples from 76 of the 86 asymptomatic persons (75 nasopharyngeal swab samples and 1 nasal swab sample) were obtained and tested by means of PCR assay. 4. On quarantine day 6, samples from all 87 quarantined asymptomatic persons (85 nasopharyngeal swab samples and 2 nasal swab samples [3 of the 6 persons who had been transferred to the hospital before February 2 had returned to the government quarantine facility]) were obtained and tested; all tested negative."
	Is there concern that the target condition as defined by the reference standard does not match the review question?	**Low**
Domain 4: Flow and timing - Risk of bias	4.1 Did all passengers receive the reference standard?	**Yes**
	4.2 Did all passengers receive the same reference standard?	**Yes**
	4.3 Were all passengers included in the analysis?	**Yes**

(Continued)

		Is there likely no or a very low risk of attrition bias?	
		4.4. Is it possible that the true disease status could have changed between the application of the index test and the reference standard? (e.g. additional infections through how the quarantine was handled?)	**Unclear**
		Comments on risk of bias	
		4. Could the passenger flow have introduced bias?	**Unclear**
Wong J 2020	Domain 1: Patient selection - A. Risk of bias	1.1 Was a consecutive or random sample of participants enrolled?	**Yes**
		1.2 Was a case-control design avoided?	**Yes**
		1.3 Did the study avoid inappropriate exclusions?	**Unclear**
		Comments	The publication provides limited information on the population arriving. The population in question comprises individuals arriving between 21.03.2020 and 24.04.2020. It can be assumed, that all passengers underwent the same procedure and received the same reference test (RT-PCR testing) but this is not clear from the publication
		1. Could the selection of participants have introduced bias?	**Unclear**
	Domain 1: Patient selection - B. Concerns regarding applicability	Describe included participants (prior testing, presentation, intended use of index test and setting):	The study reports on individuals travelling to Brunei between 21st of March and 24th of April 2020. As it is not a repatriation study, the population is close to real world travellers. However, the transferability of the travelling population in the early phase of the pandemic (those returning at times of border closure initiation in numerous countries) could be regarded as different from regular travel. The study provides limited information on the population, however.
		Is there concern that the included participants do not match the review question?	**Unclear**
	Domain 2: Index test(s) A. Risk of bias	2.1 Can we be sure that those identified in index test (true and false positive screening results) were identified by the index test (e.g. automated fever scanner) rather than any other means (e.g. self-reporting)?	**Unclear**

(Continued)

	2.2 Were the index test results interpreted without knowledge of the results of the reference standard?	**Yes**
	2.3 If a threshold was used, was it prespecified?	**Unclear**
	Comments on risk of bias	The procedures of performing and interpreting the index test is unclear; the publication provides limited information This is a hypothetical intervention; all individuals were tested with a PCR upon arrival; and those being symptomatic were reported. It cannot be ruled out that other approaches would have been conducted; if only those with symptoms were tested
	2. Could the conduct or interpretation of the index test have introduced bias?	**Unclear**
Domain 2: Index test(s) B. Concerns regarding applicability	Describe the index test and how it was conducted and interpreted:	The symptom/exposure-based screening was conducted on a large number of individuals as a program implemented on a regular basis at the airport (i.e. not just once for a repatriation flight); however the travel volume is lower than would likely be expected under real world scenarios in most situations without additional travel restrictions / border closures
	Is there concern that the index test, its conduct, or interpretation differ from the review question?	**Unclear**
Domain 3: Reference standard A. Risk of bias	3.1 Is the reference standard (the approach to identify and classify 'cases') likely to correctly classify the target condition (here active infection with SARS-CoV-2)?	**Unclear**
	3.2. Were the reference standard results interpreted without knowledge of the results of the index test?	**No**
	Comments on risk of bias	The study population received at least one PCR test upon arrival and underwent a 14 day quarantine procedure. There is limited information provided (e.g. on whether there was a second PCR-test prior to release or how the symptom-observation during the quarantine procedure was conducted. It is therefore unclear whether the reference test can be considered as sufficient
	3. Could the reference standard, its conduct, or its interpretation have introduced bias? (Leading to an un-	**Unclear**

(Continued)

			der-determination or over-determination of true findings in the reference test)	
	Domain 3: Reference standard & timing B. Concerns regarding applicability	Describe the reference standard and how it was conducted and interpreted:	The reference test for the symptom screening consisted of at least one PCR test close to the time point of arrival. As PCR testing is considered the gold standard with an assumed very high specificity, the definition of case and not case regarding SARS-CoV-2 infections can be assumed as given	
		Is there concern that the target condition as defined by the reference standard does not match the review question?	**Low**	
	Domain 4: Flow and timing - Risk of bias	4.1 Did all passengers receive the reference standard?	**Unclear**	
		4.2 Did all passengers receive the same reference standard?	**Unclear**	
		4.3 Were all passengers included in the analysis? Is there likely no or a very low risk of attrition bias?	**Unclear**	
		4.4. Is it possible that the true disease status could have changed between the application of the index test and the reference standard? (e.g. additional infections through how the quarantine was handled?)	**Unclear**	
		Comments on risk of bias	The publication provides limited information on the population arriving. The population in question comprises individuals arriving between 21.03.2020 and 24.04.2020. It can be assumed, that all passengers underwent the same procedure and received the same reference test (RT-PCR testing) but this is not clear from the publication. It is unclear, whether the 30 individuals who tested positive were identified by the PCR test upon arrival, or by another procedure.	
		4. Could the passenger flow have introduced bias?	**Unclear**	
Yamahata 2020	Domain 1: Patient selection - A. Risk of bias	1.1 Was a consecutive or random sample of participants enrolled?	**Yes**	
		1.2 Was a case-control design avoided?	**Yes**	

	1.3 Did the study avoid inappropriate exclusions?	**Yes**
	Comments	All passengers and crew members received the index test
	1. Could the selection of participants have introduced bias?	**Low**
Domain 1: Patient selection - B. Concerns regarding applicability	Describe included participants (prior testing, presentation, intended use of index test and setting):	Not a real world screening intervention. This is a situation which can be seen as analogous to a screening intervention. "On February 3, 2020, when the ship arrived at Yokohama, a quarantine was initiated. All passengers and crew underwent medical examinations. On February 5, the RT-PCR results from the throat swab for symptomatic people and their close contacts revealed that 10 of 31 individuals were positive for SARS-CoV-2. On the same day, the Japanese government decided that all passengers were to be quarantined in their cabins for 14 days [10]. Based on international guidance on infection control, the crew continued to maintain ship functions and support passengers for their food, clothing, and shelter-related needs."
	Is there concern that the included participants do not match the review question?	**High**
Domain 2: Index test(s) A. Risk of bias	2.1 Can we be sure that those identified in index test (true and false positive screening results) were identified by the index test (e.g. automated fever scanner) rather than any other means (e.g. self-reporting)?	**Unclear**
	2.2 Were the index test results interpreted without knowledge of the results of the reference standard?	**Yes**
	2.3 If a threshold was used, was it prespecified?	**Unclear**
	Comments on risk of bias	The study provides very limited information on the screening measure.
	2. Could the conduct or interpretation of the index test have introduced bias?	**Unclear**
Domain 2: Index test(s) B. Concerns regarding applicability	Describe the index test and how it was conducted and interpreted:	Not a real world screening intervention. It is a situation which can be seen as analogous to a screening intervention Likely, the awareness of the risk status of the individuals led to led

(Continued)

		to higher vigilance than would take place in a normal airport screening "On February 3, 2020, when the ship arrived at Yokohama, a quarantine was initiated. All passengers and crew underwent medical examinations. On February 5, the RT-PCR results from the throat swab for symptomatic people and their close contacts revealed that 10 of 31 individuals were positive for SARS-CoV-2. On the same day, the Japanese government decided that all passengers were to be quarantined in their cabins for 14 days [10]. Based on international guidance on infection control, the crew continued to maintain ship functions and support passengers for their food, clothing, and shelter-related needs."
	Is there concern that the index test, its conduct, or interpretation differ from the review question?	**High**
Domain 3: Reference standard A. Risk of bias	3.1 Is the reference standard (the approach to identify and classify 'cases') likely to correctly classify the target condition (here active infection with SARS-CoV-2)?	**No**
	3.2. Were the reference standard results interpreted without knowledge of the results of the index test?	**No**
	Comments on risk of bias	All passengers received a PCR test. All passengers and crew were quarantined. Those developing symptoms received another PCR test. There is a risk of cases being missed if they had no or minor symptoms during the quarantine (about 30%) of infected
	3. Could the reference standard, its conduct, or its interpretation have introduced bias? (Leading to an under-determination or over-determination of true findings in the reference test)	**High**
Domain 3: Reference standard & timing B. Concerns regarding applicability	Describe the reference standard and how it was conducted and interpreted:	"During the quarantine, RT-PCR testing of throat swabs was extended to all passengers in the following order: On February 3, 2020, when the ship arrived at Yokohama, a quarantine was initiated. All passengers and crew underwent medical examinations. On February 5, the RT-PCR results from the throat swab for symptomatic people and their close contacts revealed that 10 of 31 individuals were positive for SARS-CoV-2. On the same day, the Japanese government decided that all passengers were to be quarantined in their cabins for 14 days. Based on

(Continued)

			international guidance on infection control, the crew continued to maintain ship functions and support passengers for their food, clothing, and shelter-related needs."
		Is there concern that the target condition as defined by the reference standard does not match the review question?	**Low**
	Domain 4: Flow and timing - Risk of bias	4.1 Did all passengers receive the reference standard?	**Yes**
		4.2 Did all passengers receive the same reference standard?	**Yes**
		4.3 Were all passengers included in the analysis? Is there likely no or a very low risk of attrition bias?	**Yes**
		4.4. Is it possible that the true disease status could have changed between the application of the index test and the reference standard? (e.g. additional infections through how the quarantine was handled?)	**No**
		Comments on risk of bias	Keeping passengers for 3 weeks on the ship could have led to the development of additional cases (close contact among infected). Thus using quarantine as part of the reference test would yield an underestimation of the effectiveness of the screening measure
		4. Could the passenger flow have introduced bias?	**High**
PCR test			
Arima 2020	Domain 1: Patient selection - A. Risk of bias	1.1 Was a consecutive or random sample of participants enrolled?	**Yes**
		1.2 Was a case-control design avoided?	**Yes**
		1.3 Did the study avoid inappropriate exclusions?	**Unclear**
		Comments	All 566 passengers repatriated were screened and followed-up over the 14-day quarantine period. 3 symptomatic passengers were denied boarding. Their status is unclear. This could lead an underestimation of the effectiveness of symptom/exposure-based screening

	1. Could the selection of participants have introduced bias?	**Unclear**
Domain 1: Patient selection - B. Concerns regarding applicability	Describe included participants (prior testing, presentation, intended use of index test and setting):	The population is a highly selected sample in the early phase of the pandemic and does not represent regular air travel passengers which would travel between China and Japan
	Is there concern that the included participants do not match the review question?	**High**
Domain 2: Index test(s) A. Risk of bias	2.1 Can we be sure that those identified in index test (true and false positive screening results) were identified by the index test (e.g. automated fever scanner) rather than any other means (e.g. self-reporting)?	**Yes**
	2.2 Were the index test results interpreted without knowledge of the results of the reference standard?	**Yes**
	2.3 If a threshold was used, was it prespecified?	**Unclear**
	Comments on risk of bias	
	2. Could the conduct or interpretation of the index test have introduced bias?	**Low**
Domain 2: Index test(s) B. Concerns regarding applicability	Describe the index test and how it was conducted and interpreted:	Repatriation study, real-world generalisability should be considered. Likely, the awareness of the risk status of the individuals led to led to higher vigilance than would take place in a normal airport screening. "Day 1 entry screening by testing oropharyngeal swab samples collected from all 566 returnees at the hospitals to which they were initially transported for SARS-CoV-2 (4); all tests were based on the real-time reverse transcription PCR developed by the National Institute of Infectious Diseases"
	Is there concern that the index test, its conduct, or interpretation differ from the review question?	**High**
Domain 3: Reference standard A. Risk of bias	3.1 Is the reference standard (the approach to identify and classify 'cases') likely to correctly classify the target condition (here active infection with SARS-CoV-2)?	**Yes**

(Continued)

	3.2. Were the reference standard results interpreted without knowledge of the results of the index test?	**Unclear**
	Comments on risk of bias	Reference test was given at both entry and 14 days later at quarantine exit; there is a slight chance given the incubation and infectiousness periods that some cases were not captured, however this risk is rather low
	3. Could the reference standard, its conduct, or its interpretation have introduced bias? (Leading to an under-determination or over-determination of true findings in the reference test)	**Low**
Domain 3: Reference standard & timing B. Concerns regarding applicability	Describe the reference standard and how it was conducted and interpreted:	Exit screening for quarantined persons who remained illness-free (i.e. not already identified as a case) by collecting oropharyngeal swab samples on day 14, the end of the quarantine period.
	Is there concern that the target condition as defined by the reference standard does not match the review question?	**Low**
Domain 4: Flow and timing - Risk of bias	4.1 Did all passengers receive the reference standard?	**Yes**
	4.2 Did all passengers receive the same reference standard?	**Unclear**
	4.3 Were all passengers included in the analysis? Is there likely no or a very low risk of attrition bias?	**Yes**
	4.4. Is it possible that the true disease status could have changed between the application of the index test and the reference standard? (e.g. additional infections through how the quarantine was handled?)	**Unclear**
	Comments on risk of bias	Three out of the 12 cases were missed in the screening and testing procedure. Of those, two developed symptoms, one was in facility and one in home quarantine. The risk of the individual to have acquired the infection from other individuals in the facility is regarded as low as it is reported "facility-quarantined case-patient was in a single room; no other person from this facility acquired COVID-19 or had a positive test result at exit screening."

Cochrane
Library

Trusted evidence.
Informed decisions.
Better health.

Cochrane Database of Systematic Reviews

(Continued)

			The publication does not provide enough information to judge if there is a risk of the other two individuals could have acquired the infection after arrival in the quarantine:
		4. Could the passenger flow have introduced bias?	**Unclear**
Al-Qahtani 2020	Domain 1: Patient selection - A. Risk of bias	1.1 Was a consecutive or random sample of participants enrolled?	**Yes**
		1.2 Was a case-control design avoided?	**Yes**
		1.3 Did the study avoid inappropriate exclusions?	**Yes**
		Comments	2714 individuals arriving at the Bahrain airport underwent PCR testing and quarantine procedures
		1. Could the selection of participants have introduced bias?	**Low**
	Domain 1: Patient selection - B. Concerns regarding applicability	Describe included participants (prior testing, presentation, intended use of index test and setting):	The study reports on individuals travelling to Bahrain between 25th of February and 14th of March 2020. As it is not a repatriation study, the population is close to real world travellers. However, the transferability of the travelling population in the early phase of the pandemic (those returning at times of border closure initiation in numerous countries) could be regarded as different than regular travel
		Is there concern that the included participants do not match the review question?	**Low**
	Domain 2: Index test(s) A. Risk of bias	2.1 Can we be sure that those identified in index test (true and false positive screening results) were identified by the index test (e.g. automated fever scanner) rather than any other means (e.g. self-reporting)?	**Yes**
		2.2 Were the index test results interpreted without knowledge of the results of the reference standard?	**Yes**
		2.3 If a threshold was used, was it prespecified?	**Yes**
		Comments on risk of bias	PCR is the only means of identifying individuals; there is no indication of additional procedures (e.g. double testing for certain individuals / case definition based on reporting of individual). Due to the PCR test; the predefini-

			tion of the PCR threshold definition is implied within the use of the method
		2. Could the conduct or interpretation of the index test have introduced bias?	**Low**
	Domain 2: Index test(s) B. Concerns regarding applicability	Describe the index test and how it was conducted and interpreted:	The PCR testing was conducted as part of a regular measure which was conducted on all travellers; however the travel volume is lower than would likely be expected under real world scenarios in most situations without additional travel restrictions / border closures
		Is there concern that the index test, its conduct, or interpretation differ from the review question?	**Unclear**
	Domain 3: Reference standard A. Risk of bias	3.1 Is the reference standard (the approach to identify and classify 'cases') likely to correctly classify the target condition (here active infection with SARS-CoV-2)?	**Yes**
		3.2. Were the reference standard results interpreted without knowledge of the results of the index test?	**No**
		Comments on risk of bias	As those with a negative initial PCR test underwent quarantine and retesting prior to release from quarantine, the status of the individuals was known to the testers. But as all individuals received the test and PCR testing can be considered relatively objective, this is regarded as low risk of bias. The risk of status change is discussed in domain 4
		3. Could the reference standard, its conduct, or its interpretation have introduced bias? (Leading to an under-determination or over-determination of true findings in the reference test)	**Low**
	Domain 3: Reference standard & timing B. Concerns regarding applicability	Describe the reference standard and how it was conducted and interpreted:	Reference test was Observation during 14 day quarantine-period with retesting in the case of symptoms and second PCR-based test for all asymptomatic passengers with an initial negative first PCR-test and no symptoms during quarantine. While there is a risk of two negative PCR tests in a row without the development of symptoms, this risk is low.
		Is there concern that the target condition as defined by the reference standard does not match the review question?	**Low**

(Continued)

	Domain 4: Flow and timing - Risk of bias	4.1 Did all passengers receive the reference standard?	**Yes**
		4.2 Did all passengers receive the same reference standard?	**Yes**
		4.3 Were all passengers included in the analysis? Is there likely no or a very low risk of attrition bias?	**Yes**
		4.4. Is it possible that the true disease status could have changed between the application of the index test and the reference standard? (e.g. additional infections through how the quarantine was handled?)	**Unclear**
		Comments on risk of bias	All Individuals received the PCR test at the same point of time for each repatriation flight and by the same group of experts. There could be a low risk of between flights contamination; e.g. due to differences in taking the swab between experts. The risk of infection during the quarantine period is unclear, this is not well described. There is a risk of individuals having a first false negative PCR test. During the 14 day quarantine period, those individuals recover. At the second PCR test they are correctly tested negative. There is a low risk of the time delay between the index test and the reference test leads to missing infected individuals within the population of those with two negative PCR tests.
		4. Could the passenger flow have introduced bias?	**Unclear**
Al-Tawfiq 2020	Domain 1: Patient selection - A. Risk of bias	1.1 Was a consecutive or random sample of participants enrolled?	**Unclear**
		1.2 Was a case-control design avoided?	**Yes**
		1.3 Did the study avoid inappropriate exclusions?	**Yes**
		Comments	1928 individuals working for a large oil company of the Kingdom of Saudi Arabia and returning in the early phase of the pandemic. Likely, all individuals working for the company underwent the procedure. A number of travellers arrived through Bahrain. The study reports that those travellers arriving through Bahrain underwent a 2 weeks quarantine procedure before com-

(Continued)

		ing to Saudi Arabia. There is a risk that individuals showing symptoms may have undergone different procedures e.g. individuals with respiratory symptoms not being allowed to travel to Saudi Arabia. This could be compared to a symptom/exposure-based departure screening upon departure, with an unclear number regarding those retained.
	1. Could the selection of participants have introduced bias?	**High**
Domain 1: Patient selection - B. Concerns regarding applicability	Describe included participants (prior testing, presentation, intended use of index test and setting):	The population is characterized by their status as employees of a large oil company of Saudi Arabia. This leads to a distortion which makes the population different from regular air travel passengers (e.g. 66% males).
	Is there concern that the included participants do not match the review question?	**Unclear**
Domain 2: Index test(s) A. Risk of bias	2.1 Can we be sure that those identified in index test (true and false positive screening results) were identified by the index test (e.g. automated fever scanner) rather than any other means (e.g. self-reporting)?	**Yes**
	2.2 Were the index test results interpreted without knowledge of the results of the reference standard?	**Yes**
	2.3 If a threshold was used, was it prespecified?	**Yes**
	Comments on risk of bias	All individuals received a nasopharyngeal and oropharyngeal swab PCR test within 24h after arrival. The test was taken by trained health professionals. A bias due to the process of performing the index test is unlikely
	2. Could the conduct or interpretation of the index test have introduced bias?	**Low**
Domain 2: Index test(s) B. Concerns regarding applicability	Describe the index test and how it was conducted and interpreted:	The symptom screening was conducted on a large number of individuals as a program implemented on a regular basis at the airport (i.e. not just once for a repatriation flight); however the travel volume is lower than would likely be expected under real world scenarios in most situations without additional travel restrictions / border closures
	Is there concern that the index test, its conduct, or	**Unclear**

(Continued)

		interpretation differ from the review question?	
	Domain 3: Reference standard A. Risk of bias	3.1 Is the reference standard (the approach to identify and classify 'cases') likely to correctly classify the target condition (here active infection with SARS-CoV-2)?	**yes**
		3.2. Were the reference standard results interpreted without knowledge of the results of the index test?	**No**
		Comments on risk of bias	The reference test was a PCR based test on day 13, as well as symptom-observation during a 14 day quarantine period. As a case is defined by having a positive PCR test, it can be assumed very likely, that all individuals who tested negative were truly not infected with SARS-CoV-2 Due to the combination of symptom observation for 14 days and PCR-based testing prior to release, it can be assumed that there is a low risk of bias due to the reference test falsely classifying an individual as negative.
		3. Could the reference standard, its conduct, or its interpretation have introduced bias? (Leading to an under-determination or over-determination of true findings in the reference test)	**Low**
	Domain 3: Reference standard & timing B. Concerns regarding applicability	Describe the reference standard and how it was conducted and interpreted:	As PCR testing is considered the gold standard with an assumed very high specificity, the definition of case and not case regarding SARS-CoV-2 infections can be assumed as given
		Is there concern that the target condition as defined by the reference standard does not match the review question?	**Low**
	Domain 4: Flow and timing - Risk of bias	4.1 Did all passengers receive the reference standard?	**Yes**
		4.2 Did all passengers receive the same reference standard?	**Yes**
		4.3 Were all passengers included in the analysis? Is there likely no or a very low risk of attrition bias?	**Yes**
		4.4. Is it possible that the true disease status could have changed between	**No**

(Continued)

		the application of the index test and the reference standard? (e.g. additional infections through how the quarantine was handled?)	
		Comments on risk of bias	The study writes: "...there were strict quarantine protocols. Guests not being allowed out of the rooms except in case of emergency". There is no indication of an outbreak in the facility.
		4. Could the passenger flow have introduced bias?	**Low**
Lagier 2020	Domain 1: Patient selection - A. Risk of bias	1.1 Was a consecutive or random sample of participants enrolled?	**Yes**
		1.2 Was a case-control design avoided?	**Yes**
		1.3 Did the study avoid inappropriate exclusions?	**Unclear**
		Comments	The study reports that the quarantine procedure began in China 2-7 days prior to travelling back to France. Limited information is provided here. It is unclear, if individuals who showed symptoms upon entry into quarantine or developed symptoms during the quarantine phase were handled (e.g. did they receive a PCR test and were denied the boarding the flight?)
		1. Could the selection of participants have introduced bias?	**High**
	Domain 1: Patient selection - B. Concerns regarding applicability	Describe included participants (prior testing, presentation, intended use of index test and setting):	The population is a highly selected sample of repatriated individuals in the early phase of the pandemic and does not represent regular air travel passengers,
		Is there concern that the included participants do not match the review question?	**High**
	Domain 2: Index test(s) A. Risk of bias	2.1 Can we be sure that those identified in index test (true and false positive screening results) were identified by the index test (e.g. automated fever scanner) rather than any other means (e.g. self-reporting)?	**Yes**
		2.2 Were the index test results interpreted without knowledge of the results of the reference standard?	**Yes**
		2.3 If a threshold was used, was it prespecified?	**Unclear**

(Continued)

	Comments on risk of bias	All repatriated individuals were offered a PCR test upon arrival (within 24h). As the PCR tests are established and sufficiently defined procedures, the risk of bias due to the approach of conducting the index test is unlikely. However, the study reports that the quarantine procedure began in China 2-7 days prior to travelling back to France. Limited information is provided here. Therefore, the PCR-based screening was conducted between day 3 and 8 of quarantine procedure in China with a flight in between. It is unclear, whether this could lead to a bias regarding the ability of the PCR test to detect cases. The symptom/exposure-based screening was conducted by professional individuals at a military air force base. The procedure is not described in detail. It is likely, that a threshold for symptom-status was established (although not described in detail), as this event triggered action regarding the symptomatic individual
	2. Could the conduct or interpretation of the index test have introduced bias?	**Unclear**
Domain 2: Index test(s) B. Concerns regarding applicability	Describe the index test and how it was conducted and interpreted:	Index test is that all individuals after a period in a quarantine facility in China of 2-7 days and within 24h after arrival from China - independent of symptom status - receive a PCR test. This procedure is likely very similar to such a measure being conducted in real world airport settings
	Is there concern that the index test, its conduct, or interpretation differ from the review question?	**Low**
Domain 3: Reference standard A. Risk of bias	3.1 Is the reference standard (the approach to identify and classify 'cases') likely to correctly classify the target condition (here active infection with SARS-CoV-2)?	**yes**
	3.2. Were the reference standard results interpreted without knowledge of the results of the index test?	**No**
	Comments on risk of bias	The reference test consisted of a second PCR test (day 5 after arrival; day 7-12 after quarantine in China) as well as symptom/exposure-based observation during a 5-day (unclear) quarantine with a follow-up PCR test in the case of symptoms.

		While this procedure is likely sufficient assuming the 2-7 day quarantine procedure prior to the flight (including not described local measures), this is likely sufficient to detect cases. However, assuming that the pre-quarantine procedure did not taken place (this was already covered by the risk of bias regarding the population selection) the ability to detect cases is unclear.
	3. Could the reference standard, its conduct, or its interpretation have introduced bias? (Leading to an under-determination or over-determination of true findings in the reference test)	**Low**
Domain 3: Reference standard & timing B. Concerns regarding applicability	Describe the reference standard and how it was conducted and interpreted:	The reference test PCR testing upon arrival consisted of a 1 PCR tests as well as symptom observation during a 5 day quarantine period. As PCR-testing is considered the gold standard with an assumed very high specificity, the definition of case and not case regarding SARS-CoV-2 infections can be assumed as given.
	Is there concern that the target condition as defined by the reference standard does not match the review question?	**Low**
Domain 4: Flow and timing - Risk of bias	4.1 Did all passengers receive the reference standard?	**No**
	4.2 Did all passengers receive the same reference standard?	**No**
	4.3 Were all passengers included in the analysis? Is there likely no or a very low risk of attrition bias?	**Unclear**
	4.4. Is it possible that the true disease status could have changed between the application of the index test and the reference standard? (e.g. additional infections through how the quarantine was handled?)	**No**
	Comments on risk of bias	7/337 individuals were tested only once and 1/337 individual refused both tests. It is unclear from the study reporting, if those who refused were symptomatic upon arrival or developed symptoms during the course of the observation period (leading to a higher pre-test probability). Without additional information on the individuals who refused/were not tested; the risk of bias is unclear.

(Continued)

			Among those who were tested twice: While the quarantine measures are not described in detail, as there were in total 0 cases detected, it is unlikely that the disease status of individuals had changed due to the quarantine procedures. Due to the two PCR tests (day 1 and day 5), it can be assumed that all infections within this period (even those in decline) would have been covered.
		4. Could the passenger flow have introduced bias?	**Unclear**
Lio 2020	Domain 1: Patient selection - A. Risk of bias	1.1 Was a consecutive or random sample of participants enrolled?	**Unclear**
		1.2 Was a case-control design avoided?	**Yes**
		1.3 Did the study avoid inappropriate exclusions?	**Yes**
		Comments	All passengers on the flight were included in the study. However, one passenger with symptoms was denied boarding. The status regarding SARS-CoV-2 infection of this passenger is not reported
		1. Could the selection of participants have introduced bias?	**Unclear**
	Domain 1: Patient selection - B. Concerns regarding applicability	Describe included participants (prior testing, presentation, intended use of index test and setting):	The population is a highly selected sample of repatriated individuals in the early phase of the pandemic and does not represent regular air travel passengers,
		Is there concern that the included participants do not match the review question?	**High**
	Domain 2: Index test(s) A. Risk of bias	2.1 Can we be sure that those identified in index test (true and false positive screening results) were identified by the index test (e.g. automated fever scanner) rather than any other means (e.g. self-reporting)?	**Yes**
		2.2 Were the index test results interpreted without knowledge of the results of the reference standard?	**Yes**
		2.3 If a threshold was used, was it prespecified?	**Yes**

(Continued)

	Comments on risk of bias	The individuals underwent a 2 day quarantine with symptom observation, followed by a PCR test on day 2. As the PCR tests are established and sufficiently defined procedures, the risk of bias due to the approach of conducting the index test is unlikely
	2. Could the conduct or interpretation of the index test have introduced bias?	**Low**
Domain 2: Index test(s) B. Concerns regarding applicability	Describe the index test and how it was conducted and interpreted:	Index test is that all individuals after a period in a quarantine facility of 2 day - independent of symptom status - receive a PCR test. This procedure is likely very similar to such a measure being conducted in real world airport settings
	Is there concern that the index test, its conduct, or interpretation differ from the review question?	**Low**
Domain 3: Reference standard A. Risk of bias	3.1 Is the reference standard (the approach to identify and classify 'cases') likely to correctly classify the target condition (here active infection with SARS-CoV-2)?	**yes**
	3.2. Were the reference standard results interpreted without knowledge of the results of the index test?	**No**
	Comments on risk of bias	The reference test for the symptom based screening consisted of a 2 PCR tests at day 7 and 13 as well as symptom-observation during a 14 day quarantine phase (11-12 days after first PCR-test). It can be assumed very likely, that all individuals who tested negative were truly not infected with SARS-CoV-2
	3. Could the reference standard, its conduct, or its interpretation have introduced bias? (Leading to an under-determination or over-determination of true findings in the reference test)	**Low**
Domain 3: Reference standard & timing B. Concerns regarding applicability	Describe the reference standard and how it was conducted and interpreted:	The reference test for the symptom screening consisted of a 3 PCR tests at day 2, 7 and 13 as well as symptom observation during a 14 day quarantine period. As PCR testing is considered the gold standard with an assumed very high specificity, the definition of case and not case regarding SARS-CoV-2 infections can be assumed as given

(Continued)

		Is there concern that the target condition as defined by the reference standard does not match the review question?	**Low**
	Domain 4: Flow and timing - Risk of bias	4.1 Did all passengers receive the reference standard?	**Yes**
		4.2 Did all passengers receive the same reference standard?	**Yes**
		4.3 Were all passengers included in the analysis? Is there likely no or a very low risk of attrition bias?	**Yes**
		4.4. Is it possible that the true disease status could have changed between the application of the index test and the reference standard? (e.g. additional infections through how the quarantine was handled?)	**No**
		Comments on risk of bias	While the quarantine measures are not described in detail, as there were 0 cases detected, it is unlikely that the disease status of individuals changed due to the quarantine procedures. Due to the two PCR tests (day 2 and day 7) covering the period between arrival and the prerelease PCR-test at day 14, it can be assumed that all infections within this period (even those in decline) would have been covered.
		4. Could the passenger flow have introduced bias?	**Low**
Ng 2020	Domain 1: Patient selection - A. Risk of bias	1.1 Was a consecutive or random sample of participants enrolled?	**Yes**
		1.2 Was a case-control design avoided?	**Yes**
		1.3 Did the study avoid inappropriate exclusions?	**no**
		Comments	Here, 3 febrile passengers were removed from the pool and we do not have information on their diagnosis; thus removing infected passengers at this point would lead to an underestimate of the screening
		1. Could the selection of participants have introduced bias?	**High**

(Continued)

	Domain 1: Patient selection - B. Concerns regarding applicability	Describe included participants (prior testing, presentation, intended use of index test and setting):	Repatriation-study, transferability to the real world as a screening intervention is unclear. "We followed up on 94 persons who boarded an evacuation flight from Wuhan to Singapore on January 30, 2020."
		Is there concern that the included participants do not match the review question?	**High**
	Domain 2: Index test(s) A. Risk of bias	2.1 Can we be sure that those identified in index test (true and false positive screening results) were identified by the index test (e.g. automated fever scanner) rather than any other means (e.g. self-reporting)?	**Yes**
		2.2 Were the index test results interpreted without knowledge of the results of the reference standard?	**Yes**
		2.3 If a threshold was used, was it prespecified?	**Yes**
		Comments on risk of bias	On arrival in Singapore, the passengers underwent repeat screening for body temperature (fever was defined as a body temperature ≥38°C), and 2 persons had a fever. Co-Intervention: 1. Airport Departure fever screening: Screening for body temperature was conducted at check-in and before boarding, and 3 febrile persons were prevented from boarding (no additional information regarding the status of these 3 febrile persons was available) 2. Masks Surgical masks were provided to passengers on board the plane.
		2. Could the conduct or interpretation of the index test have introduced bias?	**Low**
	Domain 2: Index test(s) B. Concerns regarding applicability	Describe the index test and how it was conducted and interpreted:	Repatriation-study, transferability to the real world as a screening intervention is unclear. Likely, the awareness of the risk status of the individuals led to led to higher vigilance than would take place in a normal airport screening. "The 2 febrile women identified in arrival screening were transferred immediately to a hospital, and they tested positive for SARS-CoV-2 (their clinical course is described in the Supplementary Appendix)"
		Is there concern that the index test, its conduct, or	**High**

(Continued)

		interpretation differ from the review question?	
	Domain 3: Reference standard A. Risk of bias	3.1 Is the reference standard (the approach to identify and classify 'cases') likely to correctly classify the target condition (here active infection with SARS-CoV-2)?	**No**
		3.2. Were the reference standard results interpreted without knowledge of the results of the index test?	**no**
		Comments on risk of bias	While the reference test is not perfect, most passengers received at least 2 PCR tests with 3 days in between, were observed intensely for 14 days, and tested in the case of symptoms. There is a risk of participants being missed by this procedure, but likely this number is low.
		3. Could the reference standard, its conduct, or its interpretation have introduced bias? (Leading to an under-determination or over-determination of true findings in the reference test)	**High**
	Domain 3: Reference standard & timing B. Concerns regarding applicability	Describe the reference standard and how it was conducted and interpreted:	1. Quarantine and observation for 14 days and checking of symptoms and fever three times daily 2. PCR assessment of those showing symptoms; those with symptoms received only 1 test 3. On quarantine day 3, samples from 76 of the 86 asymptomatic persons (75 nasopharyngeal swab samples and 1 nasal swab sample) were obtained and tested by means of PCR assay 4. On quarantine day 6, samples from all 87 quarantined asymptomatic persons (85 nasopharyngeal swab samples and 2 nasal swab samples [3 of the 6 persons who had been transferred to the hospital before February 2 had returned to the government quarantine facility]) were obtained and tested; all tested negative.
		Is there concern that the target condition as defined by the reference standard does not match the review question?	**Low**
	Domain 4: Flow and timing - Risk of bias	4.1 Did all passengers receive the reference standard?	**Yes**
		4.2 Did all passengers receive the same	**Yes**

(Continued)

		reference standard?	
		4.3 Were all passengers included in the analysis? Is there likely no or a very low risk of attrition bias?	**Yes**
		4.4. Is it possible that the true disease status could have changed between the application of the index test and the reference standard? (e.g. additional infections through how the quarantine was handled?)	**Unclear**
		Comments on risk of bias	
		4. Could the passenger flow have introduced bias?	**Unclear**
Shaikh Abdul Karim 2020	Domain 1: Patient selection - A. Risk of bias	1.1 Was a consecutive or random sample of participants enrolled?	**Yes**
		1.2 Was a case-control design avoided?	**Yes**
		1.3 Did the study avoid inappropriate exclusions?	**Yes**
		Comments	Repatriation study of 432 individuals to Malaysia. All individuals underwent the two-step screening procedures
		1. Could the selection of participants have introduced bias?	**Low**
	Domain 1: Patient selection - B. Concerns regarding applicability	Describe included participants (prior testing, presentation, intended use of index test and setting):	The population is a highly selected sample in the early phase of the pandemic and does not represent regular air travel passengers. With a focus on the proportion of cases detected/ PPV this is likely not a source of bias; although the generalisability may be limited. Strict procedures to limit risk of infection during the flight may give higher protection against infection on the flight. The lack of these very early infections might lead to an over-estimation of sensitivity (which is very low for very early stages of infection); However it is unclear, whether this can be considered different from no protective measures.
		Is there concern that the included participants do not match the review question?	**High**
	Domain 2: Index test(s) A. Risk of bias	2.1 Can we be sure that those identified in index test (true and false positive screening results) were identi-	**Yes**

(Continued)

	fied by the index test (e.g. automated fever scanner) rather than any other means (e.g. self-reporting)?	
	2.2 Were the index test results interpreted without knowledge of the results of the reference standard?	**Yes**
	2.3 If a threshold was used, was it prespecified?	**Yes**
	Comments on risk of bias	PCR is the only means of identifying individuals; there is no indication of additional procedures to identify individuals and cases (e.g. double testing for certain individuals / case definition based on reporting of individual). Due to the PCR test; the predefinition of the PCR threshold definition is implied within the use of the method.
	2. Could the conduct or interpretation of the index test have introduced bias?	**Low**
Domain 2: Index test(s) B. Concerns regarding applicability	Describe the index test and how it was conducted and interpreted:	Repatriation study, real-world generalisability should be considered. Day 1 entry screening for symptoms was conducted by a large and specialized team; it is questionable, whether this would be possible for real world symptom/exposure-based screening for regular travellers. If the focus is the general PCR testing of all individuals, there should be limited concerns regarding the generalisability of the approach to the index test. "The reception team personnel consist of officials from NADMA, personnel from the Fire and Rescue Department as Ground Mission Commander and medical personnel from the MOH. Medical personnel from MOH comprise an emergency physician, a public health physician, nurses, assistant medical officers, a pathologist and laboratory technicians. The team utilises the AirDisaster Unit (ADU) as the base of the operations." Strict procedures to limit risk of infection during the flight may give higher protection against infection on the flight. The lack of these very early infections might lead to an over-estimation of sensitivity (which is very low for very early stages of infection); However it is unclear, whether this can be considered different from no protective measures.
	Is there concern that the index test, its conduct, or interpretation differ from the review question?	**High**

(Continued)

Domain 3: Reference standard A. Risk of bias	3.1 Is the reference standard (the approach to identify and classify 'cases') likely to correctly classify the target condition (here active infection with SARS-CoV-2)?	**Yes**
	3.2. Were the reference standard results interpreted without knowledge of the results of the index test?	**No**
	Comments on risk of bias	Reference test was observation during 14 day quarantine-period and second PCR-based test for all asymptomatic passengers with an initial negative first PCR-test. While there is a risk of two negative PCR tests in a row without the development of symptoms, this risk is low.
	3. Could the reference standard, its conduct, or its interpretation have introduced bias? (Leading to an under-determination or over-determination of true findings in the reference test)	**Low**
Domain 3: Reference standard & timing B. Concerns regarding applicability	Describe the reference standard and how it was conducted and interpreted:	Reference test was Observation during 14 day quarantine-period and second PCR-based test for all asymptomatic passengers with an initial negative first PCR-test. While there is a risk of two negative PCR tests in a row without the development of symptoms, this risk is low.
	Is there concern that the target condition as defined by the reference standard does not match the review question?	**Low**
Domain 4: Flow and timing - Risk of bias	4.1 Did all passengers receive the reference standard?	**Yes**
	4.2 Did all passengers receive the same reference standard?	**Yes**
	4.3 Were all passengers included in the analysis? Is there likely no or a very low risk of attrition bias?	**Yes**
	4.4. Is it possible that the true disease status could have changed between the application of the index test and the reference standard? (e.g. additional infections through how the quarantine was handled?)	**Unclear**

(Continued)

		Comments on risk of bias	All Individuals received the PCR test at the same point of time for each repatriation flight and by the same group of experts. There could be a low risk of between flights contamination; e.g. due to differences in taking the swab between experts. The risk of infection during the quarantine period is unclear, this is not well described. There is a risk of individuals having a first false negative PCR test. During the 14 day quarantine period, those individuals cure out the infection. At the second PCR test they are correctly tested negative. There is a low risk of the time delay between the index test and the reference test leads to missing infected individuals within the population of those with two negative PCR tests.
		4. Could the passenger flow have introduced bias?	**Unclear**

Quarantine of travellers and screening combined

Al-Qahtani 2020	Domain 1: Patient selection - A. Risk of bias	1.1 Was a consecutive or random sample of participants enrolled?	**Yes**
		1.2 Was a case-control design avoided?	**Yes**
		1.3 Did the study avoid inappropriate exclusions?	**Yes**
		Comments	2714 individuals arriving at the Bahrain airport underwent PCR testing and quarantine procedures
		1. Could the selection of participants have introduced bias?	**Low**
	Domain 1: Patient selection - B. Concerns regarding applicability	Describe included participants (prior testing, presentation, intended use of index test and setting):	The study reports on individuals travelling to Bahrain between 25 February and 14 March 2020. As it is not a repatriation study, the population is close to real world travellers. However, the transferability of the travelling population in the early phase of the pandemic (those returning at times of border closure initiation in numerous countries) could be regarded as different than regular travel
		Is there concern that the included participants do not match the review question?	**Low**
	Domain 2: Index test(s) A. Risk of bias	2.1 Can we be sure that those identified in index test (true and false positive screening results) were identified by the index test (e.g. automated fever scanner) rather than any	**Yes**

(Continued)

		other means (e.g. self-reporting)?	
	2.2 Were the index test results interpreted without knowledge of the results of the reference standard?		**Yes**
	2.3 If a threshold was used, was it prespecified?		**Unclear**
	Comments on risk of bias		Here, the index test is the symptom observation during quarantine with PCR testing on top of 1 PCR test on the day of arrival. Insufficient measures in the symptom observation and testing is likely to reduce the assumed sensitivity of the quarantine measure. 188 individuals were tested positive, 136 of which (72.3%) within arrival screening (of those 44 were symptomatic and tested positive and 92 were asymptomatic but identified through the PCR test). 27 of the remaining cases were detected through symptom-development and follow-up PCR testing. The exact procedures regarding the testing in the quarantine facility is not well described, but it is reported that systematic testing for cough, fever, and sore throat was conducted to select individuals for additional PCR testing
	2. Could the conduct or interpretation of the index test have introduced bias?		**Low**
Domain 2: Index test(s) B. Concerns regarding applicability	Describe the index test and how it was conducted and interpreted:		The index test refers to highly selective circumstances: Specialized quarantine facility with enough space for all individuals. While the transferability of these findings to similar circumstances is likely given, it is unclear, whether the results are transferable to e.g. home quarantine
	Is there concern that the index test, its conduct, or interpretation differ from the review question?		**Low**
Domain 3: Reference standard A. Risk of bias	3.1 Is the reference standard (the approach to identify and classify 'cases') likely to correctly classify the target condition (here active infection with SARS-CoV-2)?		**Yes**
	3.2. Were the reference standard results interpreted without knowledge of the results of the index test?		**No**
	Comments on risk of bias		Reference test was one additional PCR test conducted at day 14 after arrival.

(Continued)

		The risk for false negative PCR test is highest in the first days after infection with a decline in the course of the infection, likely in line with the infectiousness. It is unclear how the sensitivity and specificity of a single PCR test after 14 days of quarantine and therefore for individuals with a very long incubation period should be regarded
	3. Could the reference standard, its conduct, or its interpretation have introduced bias? (Leading to an under-determination or over-determination of true findings in the reference test)	**Unclear**
Domain 3: Reference standard & timing B. Concerns regarding applicability	Describe the reference standard and how it was conducted and interpreted:	Reference test was Observation during 14 day quarantine-period with retesting in the case of symptoms and second PCR-based test for all asymptomatic passengers with an initial negative first PCR-test and no symptoms during quarantine. While there is a risk of two negative PCR tests in a row without the development of symptoms, this risk is low.
	Is there concern that the target condition as defined by the reference standard does not match the review question?	**Low**
Domain 4: Flow and timing - Risk of bias	4.1 Did all passengers receive the reference standard?	**Yes**
	4.2 Did all passengers receive the same reference standard?	**Yes**
	4.3 Were all passengers included in the analysis? Is there likely no or a very low risk of attrition bias?	**Yes**
	4.4. Is it possible that the true disease status could have changed between the application of the index test and the reference standard? (e.g. additional infections through how the quarantine was handled?)	**Unclear**
	Comments on risk of bias	All Individuals received the PCR test at the same point of time for each repatriation flight and by the same group of experts. There could be a low risk of between flights contamination; e.g. due to differences in taking the swab between experts. The risk of infection during the quarantine period is unclear, this is not well described. As there were infections among those in quarantine, there would be a risk of infections tak-

(Continued)

			ing place in the quarantine facility. Due to the limited information, this cannot be ruled out. There is no indication of an outbreak/infections among staff in the quarantine facility Those who were tested positive were sent to specialized quarantine facilities, it is therefore unlikely that they caused infections after isolation. There is a risk of individuals having a first false negative PCR test. During the 14 day quarantine period, those individuals cure out the infection. At the second PCR test they are correctly tested negative. There is a low risk of the time delay between the index test and the reference test leads to missing infected individuals within the population of those with two negative PCR tests.
		4. Could the passenger flow have introduced bias?	**Unclear**
Arima 2020	Domain 1: Patient selection - A. Risk of bias	1.1 Was a consecutive or random sample of participants enrolled?	**Yes**
		1.2 Was a case-control design avoided?	**Yes**
		1.3 Did the study avoid inappropriate exclusions?	**Unclear**
		Comments	All 566 passengers repatriated were screened and followed-up over the 14-day quarantine period. 3 symptomatic passengers were denied boarding. Their status is unclear. This could lead to an underestimation of the effectiveness of symptom/exposure-based screening
		1. Could the selection of participants have introduced bias?	**Unclear**
	Domain 1: Patient selection - B. Concerns regarding applicability	Describe included participants (prior testing, presentation, intended use of index test and setting):	The population is a highly selected sample in the early phase of the pandemic and does not represent regular air travel passengers which would travel between China and Japan
		Is there concern that the included participants do not match the review question?	**High**
	Domain 2: Index test(s) A. Risk of bias	2.1 Can we be sure that those identified in index test (true and false positive screening results) were identified by the index test (e.g. automated fever scanner) rather than any other means (e.g. self-reporting)?	**Yes**
		2.2 Were the index test results interpreted without	**Yes**

(Continued)

		knowledge of the results of the reference standard?	
		2.3 If a threshold was used, was it prespecified?	**Unclear**
		Comments on risk of bias	
		2. Could the conduct or interpretation of the index test have introduced bias?	**Low**
	Domain 2: Index test(s) B. Concerns regarding applicability	Describe the index test and how it was conducted and interpreted:	Repatriation study, real-world generalisability should be considered. Likely, the awareness of the risk status of the individuals led to led to higher vigilance than would take place in a normal airport screening. "Day 1 entry screening by testing oropharyngeal swab samples collected from all 566 returnees at the hospitals to which they were initially transported for SARS-CoV-2 (4); all tests were based on the real-time reverse transcription PCR developed by the National Institute of Infectious Diseases."
		Is there concern that the index test, its conduct, or interpretation differ from the review question?	**High**
	Domain 3: Reference standard A. Risk of bias	3.1 Is the reference standard (the approach to identify and classify 'cases') likely to correctly classify the target condition (here active infection with SARS-CoV-2)?	**Yes**
		3.2. Were the reference standard results interpreted without knowledge of the results of the index test?	**Unclear**
		Comments on risk of bias	Reference test was given at both entry and 14 days later at quarantine exit; there is a slight risk given the incubation and infectiousness periods that some cases were not captured, however this risk of rather low
		3. Could the reference standard, its conduct, or its interpretation have introduced bias? (Leading to an under-determination or over-determination of true findings in the reference test)	**Low**
	Domain 3: Reference standard & timing B. Concerns regarding applicability	Describe the reference standard and how it was conducted and interpreted:	Exit screening for quarantined persons who remained illness-free (i.e. not already identified as a case) by collecting oropharyngeal swab samples on day 14, the end of the quarantine period.

(Continued)

		Is there concern that the target condition as defined by the reference standard does not match the review question?	**Low**
	Domain 4: Flow and timing - Risk of bias	4.1 Did all passengers receive the reference standard?	**Yes**
		4.2 Did all passengers receive the same reference standard?	**Unclear**
		4.3 Were all passengers included in the analysis? Is there likely no or a very low risk of attrition bias?	**Yes**
		4.4. Is it possible that the true disease status could have changed between the application of the index test and the reference standard? (e.g. additional infections through how the quarantine was handled?)	**Unclear**
		Comments on risk of bias	Three out of the 12 cases were missed in the screening and testing procedure. Of those, two developed symptoms, one was in facility and one in home quarantine. The risk of the individual to have acquired the infection from other individuals in the facility is regarded as low as it is reported "facility-quarantined case-patient was in a single room; no other person from this facility acquired COVID-19 or had a positive test result at exit screening." The publication does not provide enough information to judge if there is a risk of the other two individuals could have acquired the infection after arrival in the quarantine:
		4. Could the passenger flow have introduced bias?	**Unclear**
Chen J 2020	Domain 1: Patient selection - A. Risk of bias	1.1 Was a consecutive or random sample of participants enrolled?	**Yes**
		1.2 Was a case-control design avoided?	**Yes**
		1.3 Did the study avoid inappropriate exclusions?	**Yes**
		Comments	All passengers on the flight were included in the study. The crew was not included in the analysis.
		1. Could the selection of participants have introduced bias?	**Low**

(Continued)

Domain 1: Patient selection - B. Concerns regarding applicability	Describe included participants (prior testing, presentation, intended use of index test and setting):	The population is a highly selected sample of travellers where an assumed outbreak during the flight occurred
	Is there concern that the included participants do not match the review question?	**High**
Domain 2: Index test(s) A. Risk of bias	2.1 Can we be sure that those identified in index test (true and false positive screening results) were identified by the index test (e.g. automated fever scanner) rather than any other means (e.g. self-reporting)?	**Yes**
	2.2 Were the index test results interpreted without knowledge of the results of the reference standard?	**Yes**
	2.3 If a threshold was used, was it prespecified?	**Yes**
	Comments on risk of bias	Index test refers to the combination of general PCR test upon arrival and symptom-focused quarantine observation with PCR testing upon symptom development. All participants underwent symptom observation with PCR testing upon symptom development for 14 days, with passengers from Wuhan an additional 7 days. All participants received a second (reference) PCR test independent of symptom status at day 13. Participants were diagnosed using PCR test kits recommended by the Chinese CDC in accordance with protocols established by the WHO. The symptom observation during quarantine was conducted by temperature measurement twice daily and report the development of symptoms.
	2. Could the conduct or interpretation of the index test have introduced bias?	**Low**
Domain 2: Index test(s) B. Concerns regarding applicability	Describe the index test and how it was conducted and interpreted:	The index test refers to highly selective circumstances: quarantine in single rooms in a hotel with strict symptom observation and reporting of symptoms. While the transferability of these findings to similar circumstances is likely given, it is unclear, whether the results are transferable to e.g. home quarantine. All participants underwent symptom observation with PCR testing upon symptom development for 14 days, with passengers from Wuhan an additional 7 days. All participants received a second (reference) PCR test inde-

(Continued)

		pendent of symptom status at day 13. Participants were diagnosed using PCR test kits recommended by the Chinese CDC in accordance with protocols established by the WHO. The symptom observation during quarantine was conducted by temperature measurement twice daily and report the development of symptoms.
	Is there concern that the index test, its conduct, or interpretation differ from the review question?	**Low**
Domain 3: Reference standard A. Risk of bias	3.1 Is the reference standard (the approach to identify and classify 'cases') likely to correctly classify the target condition (here active infection with SARS-CoV-2)?	**yes**
	3.2. Were the reference standard results interpreted without knowledge of the results of the index test?	**No**
	Comments on risk of bias	The reference test is a PCR test prior to release (day 13 after arrival). Due to the high specificity of PCR tests, the risk of false positive PCR tests can be regarded as low. While there is some risk of false-negative results, this probability can be regarded as low due to the very low pre-test probability (this would require 2 negative PCR tests and no symptom development during the observation period).
	3. Could the reference standard, its conduct, or its interpretation have introduced bias? (Leading to an under-determination or over-determination of true findings in the reference test)	**Low**
Domain 3: Reference standard & timing B. Concerns regarding applicability	Describe the reference standard and how it was conducted and interpreted:	All participants received two PCR tests at day 0 and day 13 and symptom observation with temperature measurement twice daily with PCR testing upon symptom development was used as reference test.
	Is there concern that the target condition as defined by the reference standard does not match the review question?	**Low**
Domain 4: Flow and timing - Risk of bias	4.1 Did all passengers receive the reference standard?	**Yes**
	4.2 Did all passengers receive the same	**Yes**

(Continued)

		reference standard?	
		4.3 Were all passengers included in the analysis? Is there likely no or a very low risk of attrition bias?	**Yes**
		4.4. Is it possible that the true disease status could have changed between the application of the index test and the reference standard? (e.g. additional infections through how the quarantine was handled?)	**Unclear**
		Comments on risk of bias	Other than children with families, travellers were quarantined in single rooms in the hotel. There are limited descriptions on the quarantine regulation (e.g. on whether they were allowed to move freely). 3 individuals were identified at day 7 and day 9 after arrival. Two infected individuals were a couple with departure in Wuhan and belonging to the same tour group as most of the infected, including the likely two index participants. One individual with a symptom development at day 9 after arrival had a departure in Wuhan and did not belong to the tour group as most of the infected. It cannot be ruled out that those tested positive had developed symptoms during the quarantine process without additional information
		4. Could the passenger flow have introduced bias?	**Unclear**
Lio 2020	Domain 1: Patient selection - A. Risk of bias	1.1 Was a consecutive or random sample of participants enrolled?	**Yes**
		1.2 Was a case-control design avoided?	**Yes**
		1.3 Did the study avoid inappropriate exclusions?	**Yes**
		Comments	All passengers on the flight were included in the study. However, one passenger with symptoms was denied boarding. the status regarding SARS-CoV-2 infection of this passenger is not reported
		1. Could the selection of participants have introduced bias?	**Unclear**
	Domain 1: Patient selection - B. Concerns regarding applicability	Describe included participants (prior testing, presentation, intended use of index test and setting):	The population is a highly selected sample of repatriated individuals in the early phase of the pandemic and does not represent regular air travel passengers,

	Is there concern that the included participants do not match the review question?	**High**
Domain 2: Index test(s) A. Risk of bias	2.1 Can we be sure that those identified in index test (true and false positive screening results) were identified by the index test (e.g. automated fever scanner) rather than any other means (e.g. self-reporting)?	**N/A**
	2.2 Were the index test results interpreted without knowledge of the results of the reference standard?	**N/A**
	2.3 If a threshold was used, was it prespecified?	**Unclear**
	Comments on risk of bias	The index test (2 delayed PCR tests on day 2 and 7 as well as symptom observation during quarantine) and the reference test (1 PCR tests on day 13) did not identify any infected individuals. The process for symptom-observation during quarantine is not well described; but no individuals were identified as cases.
	2. Could the conduct or interpretation of the index test have introduced bias?	**Low**
Domain 2: Index test(s) B. Concerns regarding applicability	Describe the index test and how it was conducted and interpreted:	The screening procedure is not well described. But it was conducted not as part of a regular screening of travellers, but of a small set of repatriated individuals. However, an institutionalised quarantine measure with PCR testing would likely be close to what was conducted in this study. The transferability of this particular measure to other similar approaches needs to be considered
	Is there concern that the index test, its conduct, or interpretation differ from the review question?	**Unclear**
Domain 3: Reference standard A. Risk of bias	3.1 Is the reference standard (the approach to identify and classify 'cases') likely to correctly classify the target condition (here active infection with SARS-CoV-2)?	**yes**
	3.2. Were the reference standard results interpreted without knowledge of the results of the index test?	**Unclear**
	Comments on risk of bias	The reference test is a PCR test prior to release. Due to the high specificity of PCR tests,

(Continued)

		the risk of false positive PCR tests can be regarded as low. Due to the very low pre-test probability (2 negative PCR tests and no symptom development during the observation period) the risk of a false negative test result can be regarded as low.
	3. Could the reference standard, its conduct, or its interpretation have introduced bias? (Leading to an under-determination or over-determination of true findings in the reference test)	**Low**
Domain 3: Reference standard & timing B. Concerns regarding applicability	Describe the reference standard and how it was conducted and interpreted:	The reference test for the symptom screening consisted of a 3 PCR tests at day 2, 7 and 13 as well as symptom observation during a 14 day quarantine phase. As PCR testing is considered the gold standard with an assumed very high specificity, the definition of case and not case regarding SARS-CoV-2 infections can be assumed as given
	Is there concern that the target condition as defined by the reference standard does not match the review question?	**Low**
Domain 4: Flow and timing - Risk of bias	4.1 Did all passengers receive the reference standard?	**Yes**
	4.2 Did all passengers receive the same reference standard?	**Yes**
	4.3 Were all passengers included in the analysis? Is there likely no or a very low risk of attrition bias?	**Yes**
	4.4. Is it possible that the true disease status could have changed between the application of the index test and the reference standard? (e.g. additional infections through how the quarantine was handled?)	**No**
	Comments on risk of bias	While the quarantine measures are not described in detail, as there were in total 0 cases detected, it is unlikely that the disease status of individuals changed due to the quarantine procedures. Due to the two PCR tests (day 2 and day 7) covering the period between arrival and the prerelease PCR-test at day 14, it can be assumed that all infections within this period

(Continued)

			(even those in decline) would have been covered.
		4. Could the passenger flow have introduced bias?	**Low**
Shaikh Abdul Karim 2020	Domain 1: Patient selection - A. Risk of bias	1.1 Was a consecutive or random sample of participants enrolled?	**Yes**
		1.2 Was a case-control design avoided?	**Yes**
		1.3 Did the study avoid inappropriate exclusions?	**Yes**
		Comments	Repatriation study of 432 individuals to Malaysia. All individuals underwent the two-step screening procedures
		1. Could the selection of participants have introduced bias?	**Low**
	Domain 1: Patient selection - B. Concerns regarding applicability	Describe included participants (prior testing, presentation, intended use of index test and setting):	The population is a highly selected sample in the early phase of the pandemic and does not represent regular air travel passengers. With a focus on the proportion of cases/ PPV this is likely not a source of bias; although the generalisability may be limited. Strict procedures to limit risk of infection during the flight may give higher protection against infection on the flight. The lack of these very early infections might lead to an overestimation of proportion of cases detected (which is very low for very early stages of infection). However it is unclear, whether this can be considered different from no protective measures.
		Is there concern that the included participants do not match the review question?	**High**
	Domain 2: Index test(s) A. Risk of bias	2.1 Can we be sure that those identified in index test (true and false positive screening results) were identified by the index test (e.g. automated fever scanner) rather than any other means (e.g. self-reporting)?	**Yes**
		2.2 Were the index test results interpreted without knowledge of the results of the reference standard?	**Yes**
		2.3 If a threshold was used, was it prespecified?	**Yes**

(Continued)

		Comments on risk of bias	Here the index test is the combination of mass PCR testing + follow up quarantine. No symptom/exposure-based observation was conducted. Therefore, all individuals identified in the index test result from the PCR test. The risk of bias here is therefore equivalent to the assessment for the sensitivity of the mass RT-PCR testing
		2. Could the conduct or interpretation of the index test have introduced bias?	**Low**
Domain 2: Index test(s) B. Concerns regarding applicability		Describe the index test and how it was conducted and interpreted:	Repatriation study, real-world generalisability should be considered. However, an institutionalised quarantine measure with PCR testing would likely be close to what was conducted in this study. The transferability of this particular measure to other similar approaches needs to be considered. Strict procedures to limit risk of infection during the flight may give higher protection against infection on the flight. The lack of these very early infections might lead to an overestimation of sensitivity (which is very low for very early stages of infection); however it is unclear, whether this can be considered different from no protective measures.
		Is there concern that the index test, its conduct, or interpretation differ from the review question?	**High**
Domain 3: Reference standard A. Risk of bias		3.1 Is the reference standard (the approach to identify and classify 'cases') likely to correctly classify the target condition (here active infection with SARS-CoV-2)?	**Unclear**
		3.2. Were the reference standard results interpreted without knowledge of the results of the index test?	**No**
		Comments on risk of bias	Reference test was one additional PCR test conducted at day 14 after arrival. The risk for false negative PCR tests is highest in the first days after infection with a decline in the course of the infection, likely in line with the infectiousness. It is unclear how the sensitivity and specificity of a single PCR test after 14 days of quarantine and therefore for individuals with a very long incubation period should be regarded
		3. Could the reference standard, its conduct, or its interpretation have	**Unclear**

(Continued)

		introduced bias? (Leading to an under-determination or over-determination of true findings in the reference test)	
Domain 3: Reference standard & timing B. Concerns regarding applicability	Describe the reference standard and how it was conducted and interpreted:	Reference test was one additional PCR test conducted at day 14 after arrival. As a case is defined by having a positive PCR test and the risk for false positive PCR tests is relatively low; the reference test is in line with the review question	
	Is there concern that the target condition as defined by the reference standard does not match the review question?	**Low**	
Domain 4: Flow and timing - Risk of bias	4.1 Did all passengers receive the reference standard?	**Yes**	
	4.2 Did all passengers receive the same reference standard?	**Yes**	
	4.3 Were all passengers included in the analysis? Is there likely no or a very low risk of attrition bias?	**Yes**	
	4.4. Is it possible that the true disease status could have changed between the application of the index test and the reference standard? (e.g. additional infections through how the quarantine was handled?)	**Unclear**	
	Comments on risk of bias	All Individuals received the PCR test at the same point of time for each repatriation flight and by the same group of experts. There could be a low risk of between flights contamination; e.g. due to differences in taking the swab between experts. The risk of infection during the quarantine period is unclear, this is not well described. There is a risk of individuals having a first false negative PCR test. During the 14 day quarantine period, those individuals cure out the infection. At the second PCR test they are correctly tested negative. There is a low risk of the time delay between the index test and the reference test leads to missing infected individuals within the population of those with two negative PCR tests.	
	4. Could the passenger flow have introduced bias?	**Unclear**	

Appendix 9. Quality assessment of modelling studies

Study ID	Domain		Rating
Adekunle 2020	Model structure	1. Are the structural assumptions transparent and justified?	Moderate concerns
		1. Comments	It is not really clear what they did. They seem to have built a SEIR model of China or Wuhan; Next, a sample of the increasingly infected population flights across the world, using prepandemic flight levels until 24 January 2020. The model does not allow for infected (to a proportion) not to travel (e.g. for being hospitalized) after a while. They then report to have introduced flight restrictions, but it is not explained how this was done in the model. They report on time until community transmission, but it is not really clear what they did here.
		2. Are the structural assumptions reasonable given the overall objective, perspective and scope of the model?	Moderate concerns
		2. Comments	
	Input data	3. Are the input parameters transparent and justified?	No to minor concerns
		3. Comments	There are references provided for each parameter used in the model
		4. Are the input parameters reasonable?	Moderate concerns
		4. Comments	The infectious period of 4 days is rather short. The mortality of 1.8% is too high.
	Validation (external)	5. Has an external validation process been described?	Reported
		5. Comments	Predicted number of imported cases modelled compared with those observed
		6. Has the model been shown to be externally valid?	No to minor concerns
		6. Comments	Predicted number of imported cases is generally consistent with reported number of imported cases
	Validation (internal)	7. Has an internal validation process been described?	Not reported

(Continued)

		7. Comments	No internal validation conducted
		8. Has the model been shown to be internally valid?	Moderate concerns
		8. Comments	No internal validation conducted
	Uncertainty	9. Was there an adequate assessment of the effects of uncertainty?	No to minor concerns
		9. Comments	
	Transparency	10. Was technical documentation, in sufficient detail to allow (potentially) for replication, made available openly or under agreements that protect intellectual property?	Moderate concerns
		10. Comments	Technical documentation available, code not shared
Anderson 2020	Model structure	1. Are the structural assumptions transparent and justified?	No to minor concerns
		1. Comments	Well described adapted SEIR Model (SE2IQR) model. Travelers are introduced across all stages of infectiousness
		2. Are the structural assumptions reasonable given the overall objective, perspective and scope of the model?	Moderate concerns
		2. Comments	Does not allow for asymptomatic infections; Developing symptoms does not lead to an increased probability of behavior change (physical distancing)
	Input data	3. Are the input parameters transparent and justified?	No to minor concerns
		3. Comments	All parameters are well justified
		4. Are the input parameters reasonable?	Moderate concerns
		4. Comments	The incubation period of 1.2 days and the latency period of 0.2 days seems very short; and do not match the cited references. The model assumes that - across the time of infectiousness - only 1/3 of infected are put into quarantine. This figure seems relatively low (but unclear)
	Validation (external)	5. Has an external validation process been described?	Reported

(Continued)

		5. Comments	Figure S2: Model was fitted against historical model in the 12 regions
		6. Has the model been shown to be externally valid?	No to minor concerns
		6. Comments	Figure S2: Model was fitted against historical model in the 12 regions
	Validation (internal)	7. Has an internal validation process been described?	Not reported
		7. Comments	No internal validation conducted
		8. Has the model been shown to be internally valid?	Moderate concerns
		8. Comments	No internal validation conducted
	Uncertainty	9. Was there an adequate assessment of the effects of uncertainty?	Moderate concerns
		9. Comments	The model explores the implications of 1 additional case (infected traveller) for an increase in contact rates by factors of 1-2. Other parameters are not varied in the model
	Transparency	10. Was technical documentation, in sufficient detail to allow (potentially) for replication, made available openly or under agreements that protect intellectual property?	No to minor concerns
		10. Comments	Code to reproduce analysis is provided
Anzai 2020	Model structure	1. Are the structural assumptions transparent and justified?	Moderate concerns
		1. Comments	Part 1: Reduced Number of exported cases: Observed cases versus expected number (according to the model) for day 58 to day 67 following the lockdown of Wuhan Part 2: Reduced probability of a major overseas epidemic Part 3. time Delay to a Major Epidemic Gained from the Reduction in Travel Volume The structure of the model is simple; but not clearly described and not very well justified
		2. Are the structural assumptions reasonable given the overall objective, perspective and scope of the model?	Major concerns
		2. Comments	Part 1 and also part 2 and 3 of the model (as these are based on the output of part 1) rely on the num-

(Continued)

			ber of cases averted by the travel restrictions for the period day 58 to day 67: The model is a simple counterfactual model. The interruption is the lockdown of Wuhan, which took place on day 58. The model assumes, that there would have been an exponential growth in cases outside China, if the lockdown would not have been initiated, assuming that the exponential growth rate would have remained constant. This is compared against the observed number of new cases (imported and local) The difference is attributed to the travel restrictions; primarily attributed to the cordon sanitaire around Wuhan.
Input data	3. Are the input parameters transparent and justified?	No to minor concerns	
	3. Comments	As the assumed model is very simple, the input parameters used are described and justified	
	4. Are the input parameters reasonable?	Major concerns	
	4. Comments	The main input parameter is the exponential growth rate (r). It assumes, that r would remain constant for the period day 58 to day 67 and therefore does not assume any other potential confounders in China (e.g. hygiene, social distancing etc.) and outside China (travel related measures, quarantine of infected would have taken place The model relies on diagnosed cases outside china (using RT-PCR). There is a high likelihood of infected cases missed in this approach The assumptions on the R0 and the dispersion parameters k are justified and reasonable; The assumption of detected cases outside China (day 0 - 67) however is not, as we cannot be sure whether these figures are reliable; Additional testing (e.g. of repatriated passengers under quarantine) could have inflated this figure, while insufficient testing or inadequate tests could have led to a significant underestimation of h(t)	
Validation (external)	5. Has an external validation process been described?	Reported	
	5. Comments	Yes, the model has been fitted to data points for day 0-58 outside china	
	6. Has the model been shown to be externally valid?	No to minor concerns	
	6. Comments	Data points up to day 58 was used to fit the Poisson model; and yes, it does fit the model. However, these data points are unreliable	

(Continued)

	Validation (internal)	7. Has an internal validation process been described?	Not reported
		7. Comments	No internal validation conducted
		8. Has the model been shown to be internally valid?	Moderate concerns
		8. Comments	No internal validation conducted
	Uncertainty	9. Was there an adequate assessment of the effects of uncertainty?	Moderate concerns
		9. Comments	The model provides a range of sensitivity analysis, e.g. regarding r or on contact tracing. But likely additional sensitivity analysis (e.g. on underestimated cases) would have been helpful
	Transparency	10. Was technical documentation, in sufficient detail to allow (potentially) for replication, made available openly or under agreements that protect intellectual property?	Moderate concerns
		10. Comments	Code not available; formulas, input parameters, and program used for calculating the model are described
Ashcroft 2020	Model structure	1. Are the structural assumptions transparent and justified?	No to minor concerns
		1. Comments	Structure is transparently described and justified based on empirically estimated transmission parameters from the literature.
		2. Are the structural assumptions reasonable given the overall objective, perspective and scope of the model?	No to minor concerns
		2. Comments	Within the scope of the model the assumptions and structural decisions are reasonable.
	Input data	3. Are the input parameters transparent and justified?	No to minor concerns
		3. Comments	Parameters are transparently described based on the literature. Concept of the paper is to compare different possible parameter values of transmission parameters which are displayed in graphs.
		4. Are the input parameters reasonable?	No to minor concerns
		4. Comments	

(Continued)

	Validation (external)	5. Has an external validation process been described?	Not reported
		5. Comments	No external validation
		6. Has the model been shown to be externally valid?	Moderate concerns
		6. Comments	No external validation
	Validation (internal)	7. Has an internal validation process been described?	Not reported
		7. Comments	No internal validation
		8. Has the model been shown to be internally valid?	Moderate concerns
		8. Comments	Shiny app is provided which can be used to test face validity.
	Uncertainty	9. Was there an adequate assessment of the effects of uncertainty?	Major concerns
		9. Comments	Uncertainty was not assessed.
	Transparency	10. Was technical documentation, in sufficient detail to allow (potentially) for replication, made available openly or under agreements that protect intellectual property?	No to minor concerns
		10. Comments	Transparent reporting and provision of R-Shiny app.
Banholzer 2020	Model structure	1. Are the structural assumptions transparent and justified?	No to minor concerns
		1. Comments	
		2. Are the structural assumptions reasonable given the overall objective, perspective and scope of the model?	Major concerns
		2. Comments	Model assumes equal intervention effectiveness across countries and that interventions have a unique effectiveness independent of time and place they are implemented in, which is not reasonable, given that a travel ban of different countries (different levels of disease importation pressure) and at different points of time is likely to have a different effect (as the number of imported cases changes with the course of the pandemic around them).

(Continued)

			It implicitly assumes that all changes in case numbers are due to the intervention (tried to handle with sensitivity analysis)
Input data	3. Are the input parameters transparent and justified?		No to minor concerns
	3. Comments		No disease transmission parameters used due to type of model. The model uses data points of cases in the respective countries
	4. Are the input parameters reasonable?		No to minor concerns
	4. Comments		No disease transmission parameters used due to type of model. The model uses data points of cases in the respective countries
Validation (external)	5. Has an external validation process been described?		Reported
	5. Comments		Cross-validation to estimate influence of single countries.
	6. Has the model been shown to be externally valid?		No to minor concerns
	6. Comments		
Validation (internal)	7. Has an internal validation process been described?		Not reported
	7. Comments		No internal validation
	8. Has the model been shown to be internally valid?		Moderate concerns
	8. Comments		No internal validation
Uncertainty	9. Was there an adequate assessment of the effects of uncertainty?		No to minor concerns
	9. Comments		
Transparency	10. Was technical documentation, in sufficient detail to allow (potentially) for replication, made available openly or under agreements that protect intellectual property?		No to minor concerns
	10. Comments		Code published and extensive supplementary material available.

(Continued)

Bays 2020	Model structure	1. Are the structural assumptions transparent and justified?	No to minor concerns
		1. Comments	Assumptions are transparently reported.
		2. Are the structural assumptions reasonable given the overall objective, perspective and scope of the model?	Major concerns
		2. Comments	Very simplistic model. Assumptions restrict relevance of the model to a large degree. Key mechanisms like false-positive screening results or transmission among travellers are subject to unrealistic assumptions.
	Input data	3. Are the input parameters transparent and justified?	No to minor concerns
		3. Comments	Parameters are transparently described and corresponding sources are cited.
		4. Are the input parameters reasonable?	Moderate concerns
		4. Comments	Information about flight time distributions only based on assumptions.
	Validation (external)	5. Has an external validation process been described?	Not reported
		5. Comments	No external validation
		6. Has the model been shown to be externally valid?	Moderate concerns
		6. Comments	No external validation
	Validation (internal)	7. Has an internal validation process been described?	Not reported
		7. Comments	No internal validation
		8. Has the model been shown to be internally valid?	Moderate concerns
		8. Comments	Python code is available.
	Uncertainty	9. Was there an adequate assessment of the effects of uncertainty?	Major concerns
		9. Comments	Only some scenario analyses but no formal uncertainty.
	Transparency	10. Was technical documentation, in sufficient detail to allow (potentially) for replication,	No to minor concerns

(Continued)

			made available openly or under agreements that protect intellectual property?	
		10. Comments	Python code is shared and transparent reporting.
Binny 2020	Model structure	1. Are the structural assumptions transparent and justified?	No to minor concerns
		1. Comments	Model assumptions and equations are described in the text, and are justified with the current scientific literature.
		2. Are the structural assumptions reasonable given the overall objective, perspective and scope of the model?	No to minor concerns
		2. Comments	
	Input data	3. Are the input parameters transparent and justified?	No to minor concerns
		3. Comments	Input parameters and data are based on the available literature and country-specific sources.
		4. Are the input parameters reasonable?	No to minor concerns
		4. Comments	
	Validation (external)	5. Has an external validation process been described?	Not reported
		5. Comments	
		6. Has the model been shown to be externally valid?	Moderate concerns
		6. Comments	No external validation procedures reported but model fits well to observed data (external publication).
	Validation (internal)	7. Has an internal validation process been described?	Not reported
		7. Comments	No internal validation reported.
		8. Has the model been shown to be internally valid?	Moderate concerns
		8. Comments	No internal validation
	Uncertainty	9. Was there an adequate assessment of the effects of uncertainty?	No to minor concerns

(Continued)

		9. Comments	Uncertainty is assessed using 5000 realisations of the stochastic model and sensitivity analyses of assumptions.
	Transparency	10. Was technical documentation, in sufficient detail to allow (potentially) for replication, made available openly or under agreements that protect intellectual property?	Moderate concerns
		10. Comments	Model only described in another external preprint publication and no code available.
Boldog 2020	Model structure	1. Are the structural assumptions transparent and justified?	No to minor concerns
		1. Comments	clear & concise description
		2. Are the structural assumptions reasonable given the overall objective, perspective and scope of the model?	No to minor concerns
		2. Comments	
	Input data	3. Are the input parameters transparent and justified?	No to minor concerns
		3. Comments	Comprehensive description and justification of input parameters
		4. Are the input parameters reasonable?	Moderate concerns
		4. Comments	Input parameters seem to be reasonable; based on early data; change in R0 in China not adequately represented in the model
	Validation (external)	5. Has an external validation process been described?	Not reported
		5. Comments	No external validation was conducted
		6. Has the model been shown to be externally valid?	Moderate concerns
		6. Comments	No external validation
	Validation (internal)	7. Has an internal validation process been described?	Not reported
		7. Comments	No internal validation conducted
		8. Has the model been shown to be internally valid?	Moderate concerns

(Continued)

		8. Comments	No internal validation conducted
	Uncertainty	9. Was there an adequate assessment of the effects of uncertainty?	Moderate concerns
		9. Comments	the impact of variation in 3-4 key model parameters on disease outbreak risk is assessed
	Transparency	10. Was technical documentation, in sufficient detail to allow (potentially) for replication, made available openly or under agreements that protect intellectual property?	No to minor concerns
		10. Comments	Comprehensive appendix, source code for the model is available on Github
Chen Y-H 2020	Model structure	1. Are the structural assumptions transparent and justified?	No to minor concerns
		1. Comments	Assumptions are transparently reported and the appropriate literature is cited and discussed.
		2. Are the structural assumptions reasonable given the overall objective, perspective and scope of the model?	Moderate concerns
		2. Comments	Within the scope of the SEIR model the assumptions and structural decisions are reasonable. The intervention is operationalised as changes in influx of infected individuals. The main issue is with the infectiousness: those who are infected without symptoms and those who are infected and show symptoms have the same transmission probability. e.g. individuals with symptoms do not change their behavior until they become hospitalized and those who show do not show symptoms do not get a lower transmission factor
	Input data	3. Are the input parameters transparent and justified?	No to minor concerns
		3. Comments	Parameters are transparently described and corresponding sources are cited. Some parameters are varied according to potential policy scenarios.
		4. Are the input parameters reasonable?	Moderate concerns
		4. Comments	The travel restriction measure is a hypothetical intervention: the study explores the implications for the local course of the pandemic if 10 versus 5 versus 1 infected individual per day arrive (in the sense of a 50% or 90% reduction). It does not ad-

![Cochrane Library logo] Cochrane
Library

Trusted evidence.
Informed decisions.
Better health.

Cochrane Database of Systematic Reviews

(Continued)

			dress how this would take place or explores this measure in depth.
	Validation (external)	5. Has an external validation process been described?	Not reported
		5. Comments	No external validation
		6. Has the model been shown to be externally valid?	Moderate concerns
		6. Comments	No external validation
	Validation (internal)	7. Has an internal validation process been described?	Not reported
		7. Comments	No internal validation reported
		8. Has the model been shown to be internally valid?	Moderate concerns
		8. Comments	Process for internal validation not reported
	Uncertainty	9. Was there an adequate assessment of the effects of uncertainty?	Major concerns
		9. Comments	Uncertainty for the individual parameters not reported. For the scenario with the different level of external infected influx (the travel restriction measure) no alternative scenarios or parameters are explored
	Transparency	10. Was technical documentation, in sufficient detail to allow (potentially) for replication, made available openly or under agreements that protect intellectual property?	Moderate concerns
		10. Comments	Code not reported, but inputs and assumptions reported sufficiently and should allow for replication
Chen T 2020	Model structure	1. Are the structural assumptions transparent and justified?	No to minor concerns
		1. Comments	Structure is a relatively simple SEIR model; structure of model and relational assumption are adequately described; minor concerns as not the full model is provided in graphical form.
		2. Are the structural assumptions reasonable given the overall objective, perspective and scope of the model?	No to minor concerns
		2. Comments	Simple model with reasonable assumptions on the topic at hand

(Continued)

	Input data	3. Are the input parameters transparent and justified?	No to minor concerns
		3. Comments	Parameters are transparently described based on the literature. Estimation of the importation of cases is reported in more detail in the appendix
		4. Are the input parameters reasonable?	No to minor concerns
		4. Comments	Operationalization of the measures is reasonable. Input parameters are sensible. Others (Rt) were derived from real world data Estimation of importations is reasonable
	Validation (external)	5. Has an external validation process been described?	Reported
		5. Comments	Model was validated against the confirmed cases in the two countries
		6. Has the model been shown to be externally valid?	Moderate concerns
		6. Comments	The simulated numbers of newly infected cases derived from the model based on estimated time-varying Rt, actual intensity of entry restrictions and quarantine policy after adjusting a time lag between infection and reporting fit well with the official case figures. Likely, this should be regarded as a dependent validation.
	Validation (internal)	7. Has an internal validation process been described?	Not reported
		7. Comments	No internal validation reported
		8. Has the model been shown to be internally valid?	Moderate concerns
		8. Comments	Process for internal validation not reported
	Uncertainty	9. Was there an adequate assessment of the effects of uncertainty?	No to minor concerns
		9. Comments	Authors conducted a Markov Chain Monte Carlo simulation by sampling from empirical distributions of these parameters to adjust for parameter uncertainty. Parameters were derived from the literature. This was the case for the proportion of presymptomatic travellers, travel probability, incubation period and serial interval.
	Transparency	10. Was technical documentation, in sufficient detail to allow (potentially) for replication,	Moderate concerns

(Continued)

			made available openly or under agreements that protect intellectual property?	

			10. Comments	Code not reported, but inputs and assumptions reported sufficiently and should allow for replication
Chinazzi 2020	Model structure	1. Are the structural assumptions transparent and justified?	No to minor concerns
		1. Comments	
		2. Are the structural assumptions reasonable given the overall objective, perspective and scope of the model?	No to minor concerns
		2. Comments	
	Input data	3. Are the input parameters transparent and justified?	No to minor concerns
		3. Comments	
		4. Are the input parameters reasonable?	Moderate concerns
		4. Comments	Based on early data on transmission parameters; broad sensitivity analysis.
	Validation (external)	5. Has an external validation process been described?	Reported
		5. Comments	Limited approach to validate model projections against reported cases.
		6. Has the model been shown to be externally valid?	Moderate concerns
		6. Comments	Limited approach to validate model projections against reported cases.
	Validation (internal)	7. Has an internal validation process been described?	Not reported
		7. Comments	No internal validation described
		8. Has the model been shown to be internally valid?	Moderate concerns
		8. Comments	Established model, that is potentially validated, but this is not reported
	Uncertainty	9. Was there an adequate assessment of the effects of uncertainty?	No to minor concerns
		9. Comments	

(Continued)

	Transparency	10. Was technical documentation, in sufficient detail to allow (potentially) for replication, made available openly or under agreements that protect intellectual property?	Moderate concerns
		10. Comments	Technical documentation in referenced methods paper; code not available
Clifford 2020a	Model structure	1. Are the structural assumptions transparent and justified?	No to minor concerns
		1. Comments	
		2. Are the structural assumptions reasonable given the overall objective, perspective and scope of the model?	No to minor concerns
		2. Comments	
	Input data	3. Are the input parameters transparent and justified?	No to minor concerns
		3. Comments	
		4. Are the input parameters reasonable?	Major concerns
		4. Comments	Assume an exponential growth rate of r = 0.1 corresponding to an epidemic doubling time of 7.4 days based on data on early transmission in Wuhan. Also they assume a fixed travel time of 12 hours; The assumed sensitivity, cited from Quilty 2020 is questionably high. Quility data has major concerns regarding input parameters
	Validation (external)	5. Has an external validation process been described?	Not reported
		5. Comments	No external validation
		6. Has the model been shown to be externally valid?	Moderate concerns
		6. Comments	No external validation
	Validation (internal)	7. Has an internal validation process been described?	Not reported
		7. Comments	No internal validation
		8. Has the model been shown to be internally valid?	Moderate concerns

(Continued)

		8. Comments	No internal validation
	Uncertainty	9. Was there an adequate assessment of the effects of uncertainty?	No to minor concerns
		9. Comments	The authors provide a shiny app which allows readers to assess the sensitivity of results to many parameter assumptions
	Transparency	10. Was technical documentation, in sufficient detail to allow (potentially) for replication, made available openly or under agreements that protect intellectual property?	No to minor concerns
		10. Comments	Code available on GitHub
Clifford 2020b	Model structure	1. Are the structural assumptions transparent and justified?	No to minor concerns
		1. Comments	Assumptions are transparently reported and the appropriate literature is cited and discussed.
		2. Are the structural assumptions reasonable given the overall objective, perspective and scope of the model?	No to minor concerns
		2. Comments	Within the scope of the model the assumptions and structural decisions are reasonable.
	Input data	3. Are the input parameters transparent and justified?	No to minor concerns
		3. Comments	Parameters are transparently described and corresponding sources are cited. Some parameters are varied according to potential policy scenarios.
		4. Are the input parameters reasonable?	No to minor concerns
		4. Comments	
	Validation (external)	5. Has an external validation process been described?	Not reported
		5. Comments	No external validation
		6. Has the model been shown to be externally valid?	Moderate concerns
		6. Comments	No external validation
	Validation (internal)	7. Has an internal validation process been described?	Not reported

(Continued)

		7. Comments	No internal validation
		8. Has the model been shown to be internally valid?	Moderate concerns
		8. Comments	R model code is available.
	Uncertainty	9. Was there an adequate assessment of the effects of uncertainty?	No to minor concerns
		9. Comments	Uncertainty is assessed with bootstrapped confidence intervals and multiple scenarios are analysed and discussed.
	Transparency	10. Was technical documentation, in sufficient detail to allow (potentially) for replication, made available openly or under agreements that protect intellectual property?	No to minor concerns
		10. Comments	Code for analyses shared in Github repository
Costantino 2020	Model structure	1. Are the structural assumptions transparent and justified?	No to minor concerns
		1. Comments	
		2. Are the structural assumptions reasonable given the overall objective, perspective and scope of the model?	Major concerns
		2. Comments	The model does not consider other influencing factors e.g. the cordon around Wuhan as a reason for the difference and attributes it to the Australian travel ban.
	Input data	3. Are the input parameters transparent and justified?	No to minor concerns
		3. Comments	
		4. Are the input parameters reasonable?	Moderate concerns
		4. Comments	Input parameters seem to be reasonable The mortality is relatively high; The assumptions on detected and isolated cases relatively low
	Validation (external)	5. Has an external validation process been described?	Reported
		5. Comments	The estimated incidence data in China was compared to the observed incidence data; no other external validation was done

(Continued)

		6. Has the model been shown to be externally valid?	Moderate concerns
		6. Comments	Limited approach to validate model projections against reported cases.
	Validation (internal)	7. Has an internal validation process been described?	Not reported
		7. Comments	No internal validation
		8. Has the model been shown to be internally valid?	Moderate concerns
		8. Comments	No internal validation
	Uncertainty	9. Was there an adequate assessment of the effects of uncertainty?	No to minor concerns
		9. Comments	A broad range of SE with alternative assumptions of different input parameters was done
	Transparency	10. Was technical documentation, in sufficient detail to allow (potentially) for replication, made available openly or under agreements that protect intellectual property?	Moderate concerns
		10. Comments	Supplementary material with further details of the model is supposed to available, but can't be opened; code is not available
Davis 2020	Model structure	1. Are the structural assumptions transparent and justified?	No to minor concerns
		1. Comments	Model assumptions are described and extensive information is available in the supplementary material and on the project website.
		2. Are the structural assumptions reasonable given the overall objective, perspective and scope of the model?	No to minor concerns
		2. Comments	Some structural assumptions (e.g., homogenous travel probabilities) may not hold but generally minor concerns.
	Input data	3. Are the input parameters transparent and justified?	Moderate concerns
		3. Comments	Not all input parameters are transparently described but the relevant literature is cited and calibration to observed data is described.

		4. Are the input parameters reasonable?	Moderate concerns
		4. Comments	
	Validation (external)	5. Has an external validation process been described?	Reported
		5. Comments	Pearson correlation in Figure-1B
		6. Has the model been shown to be externally valid?	Moderate concerns
		6. Comments	Very limited external validation performed.
	Validation (internal)	7. Has an internal validation process been described?	Not reported
		7. Comments	No internal validation reported.
		8. Has the model been shown to be internally valid?	Moderate concerns
		8. Comments	Model is continuously developed and provided as a stand-alone software.
	Uncertainty	9. Was there an adequate assessment of the effects of uncertainty?	No to minor concerns
		9. Comments	Uncertainty is assessed with confidence intervals and posterior probability distributions and sensitivity analyses in supplementary material.
	Transparency	10. Was technical documentation, in sufficient detail to allow (potentially) for replication, made available openly or under agreements that protect intellectual property?	Moderate concerns
		10. Comments	Model is available as free, stand-alone software but no accessible source code.
Deeb 2020	Model structure	1. Are the structural assumptions transparent and justified?	No to minor concerns
		1. Comments	Model assumptions and equations are described in the text, and are justified with the current scientific literature.
		2. Are the structural assumptions reasonable given the overall objective, perspective and scope of the model?	No to minor concerns

(Continued)

	2. Comments	With the information provided, the structural assumptions underlying the SEIR model appear to be valid.
Input data	3. Are the input parameters transparent and justified?	No to minor concerns
	3. Comments	Input parameters for the transmission-related aspects underlying the model, as well as the running COVID-19 cases and travel data are based on reported data.
	4. Are the input parameters reasonable?	No to minor concerns
	4. Comments	Input parameters seem to be appropriate as far as it can be assessed.
Validation (external)	5. Has an external validation process been described?	Reported
	5. Comments	Figure-2 compares the cumulative number of infections predicted by the SEIR model with the observed data for the 130 retrospective days of the study period.
	6. Has the model been shown to be externally valid?	No to minor concerns
	6. Comments	Only dependent validation carried out; however, Figure-2 shows that the model predicted the observed cases over 130 days well.
Validation (internal)	7. Has an internal validation process been described?	Not reported
	7. Comments	No internal validation reported
	8. Has the model been shown to be internally valid?	Moderate concerns
	8. Comments	No internal validation
Uncertainty	9. Was there an adequate assessment of the effects of uncertainty?	Major concerns
	9. Comments	Additional analyses assess the impact of a variety of R values which represent varying levels of intervention; however, no analyses exploring how assumptions (such as starting values for R, E or I) may have influenced results.
Transparency	10. Was technical documentation, in sufficient detail to allow (potentially) for replication, made available openly or under	Moderate concerns

(Continued)

		agreements that protect intellectual property?	
		10. Comments	Model formulas are rather explicit but neither the code nor the data are provided. Authors write that data are available upon request.
Dickens 2020	Model structure	1. Are the structural assumptions transparent and justified?	Moderate concerns
		1. Comments	The simulation of the passengers based on general epidemiological parameters is sensible in general; However, the assumptions regarding the estimation of number of infected travellers based on overall country parameter without taking further country and population characteristics into account is likely to distort the data
		2. Are the structural assumptions reasonable given the overall objective, perspective and scope of the model?	Major concerns
		2. Comments	The model assumes the same time for illness onset to death for all countries included in the model. Given the heterogeneity of the countries (ranging from LIC to HIC), it is not reasonable. The estimation of infection burden within a country is based on the official case and death figures of the countries. Due to the different testing strategies in place, there is considerable heterogeneity in the relation of these figures in regard to the actual number of infections and COVID-19 related deaths. The approach to estimate the number of those who are "mildly symptomatic and able to Travel at T" does not adequately capture this number and is likely highly distorted due to the general testing strategy and approach within the countries.
	Input data	3. Are the input parameters transparent and justified?	No to minor concerns
		3. Comments	Parameters are transparently described based on the literature.
		4. Are the input parameters reasonable?	Moderate concerns
		4. Comments	The parameters in general have a reasonable foundation and are well justified in research. However, the generalisability of the parameters for CFR, IFR as well as time to death and time to hospitalisation for all countries is not reasonable. Regarding the PCR testing, the jump from 0% to 85% sensitivity at day two is likely inadequate; a less dichotomous development of the sensitivity is more likely.

(Continued)

	Validation (external)	5. Has an external validation process been described?	Not reported
		5. Comments	No external validation
		6. Has the model been shown to be externally valid?	Moderate concerns
		6. Comments	No external validation
	Validation (internal)	7. Has an internal validation process been described?	Not reported
		7. Comments	No internal validation reported
		8. Has the model been shown to be internally valid?	Moderate concerns
		8. Comments	Process for internal validation not reported
	Uncertainty	9. Was there an adequate assessment of the effects of uncertainty?	Major concerns
		9. Comments	Uncertainty was not assessed systematically
	Transparency	10. Was technical documentation, in sufficient detail to allow (potentially) for replication, made available openly or under agreements that protect intellectual property?	Moderate concerns
		10. Comments	Code not reported, but inputs and assumptions reported sufficiently and should allow for replication
Gostic 2020	Model structure	1. Are the structural assumptions transparent and justified?	No to minor concerns
		1. Comments	
		2. Are the structural assumptions reasonable given the overall objective, perspective and scope of the model?	Moderate concerns
		2. Comments	Model assumes two modes of screening (self-reporting / symptom screening) at arrival and departure, assuming the sensitivity of the scanners, the awareness about exposure, and the willingness to report symptoms. They furthermore assume that patients after a time period are being isolated/hospitalized in the departure country. The structural assumptions are reasonable, however the assumption, that infected are removed from the population and not able to travel seems quite arbitrary and - given that 30 are asymptomatic - the

(Continued)

			likelihood of 100% quarantine after 3-7 days is not reasonable
Input data	3. Are the input parameters transparent and justified?		No to minor concerns
	3. Comments		Authors provide references for most parameters. The parameters of awareness are assumed.
	4. Are the input parameters reasonable?		Moderate concerns
	4. Comments		The travel time of 24h hours seems to be rather long The assumption, the assumption, that infected are removed from the population and not able to ravel seems quite arbitrary and - given that 30 are asymptomatic - the likelihood of 100% quarantine after 3-7 days is not reasonable
Validation (external)	5. Has an external validation process been described?		Not reported
	5. Comments		No formal external validation conducted; however the results of the study are discussed in the light of findings of empirical studies coming to similar conclusions.
	6. Has the model been shown to be externally valid?		Moderate concerns
	6. Comments		No formal external validation conducted; However the results of the study are discussed in the light of findings of empirical studies coming to similar conclusions. Our conclusion that screening would detect no more than half of infected travellers in a growing epidemic is consistent with recent studies that have compared country-specific air travel volumes with detected case counts to estimate that roughly two thirds of imported cases remain undetected (Niehus et al., 2020; Bhatia et al., 2020). Furthermore, the finding that the majority of cases missed by screening are fundamentally undetectable is consistent with observed outcomes so far. Analyzing a line list of 290 cases imported into various countries (Dorigatti et al., 2020), we found that symptom onset occurred after the date of inbound travel for 72% (75/104) of cases for whom both dates were available, and a further 14% (15/104) had symptom onset on the date of travel. Even among passengers of repatriation flights, or quarantined on a cruise ship off the coast of Japan (who are all demonstrably at high risk), numerous cases have been undetectable in symptom screening, but have still tested positive for SARS-CoV-2 by PCR (Dorigatti et al., 2020; Hoehl et al., 2020; Japan Ministry of Health, Labor and Welfare, 2020; Nishiura et al., 2020; Hu et al., 2020). The onset of viral shedding prior to the onset of symptoms, or in cases

(Continued)

			that remain asymptomatic, is a classic factor that makes infectious disease outbreaks difficult to control (Fraser et al., 2004).
	Validation (internal)	7. Has an internal validation process been described?	Not reported
		7. Comments	No internal validation
		8. Has the model been shown to be internally valid?	Moderate concerns
		8. Comments	No internal validation
	Uncertainty	9. Was there an adequate assessment of the effects of uncertainty?	No to minor concerns
		9. Comments	Various sensitivity analysis were conducted to assess various parameters; the main parameter: number of asymptomatic cases is estimated to be 5%, 25%, and 50%
	Transparency	10. Was technical documentation, in sufficient detail to allow (potentially) for replication, made available openly or under agreements that protect intellectual property?	No to minor concerns
		10. Comments	Code available
Grannell 2020	Model structure	1. Are the structural assumptions transparent and justified?	No to minor concerns
		1. Comments	Model setup and assumptions are elaborated and justified in the text.
		2. Are the structural assumptions reasonable given the overall objective, perspective and scope of the model?	Major concerns
		2. Comments	In their two region SEIR model, the authors account for the interaction between the population living on either side of the border strip between Ireland and Northern Ireland. Within each country, they make an assumption of homogeneous mixing. It is highly unlikely that infected individuals living on the border strip between Ireland and Northern Ireland have a constant contact frequency with susceptibles from all regions from their countries. In this vein, in the case of border interactions, it is reasonable to assume that the number of susceptibles in Northern Ireland will primarily affect the Irish population living on the border strip and only affect the rest of Ireland to the extent that infected individuals living on the border strip of Ireland interact with susceptibles from other regions. There-

(Continued)

			fore, while the assumption of homogeneous mixing is a standard assumption in SEIR models, it will likely lead to an overestimation of the effects of border interactions in the context of a two region SEIR model.
Input data	3. Are the input parameters transparent and justified?	Moderate concerns	
	3. Comments	It is unclear how the authors determine some of the input parameters. For instance, they assume that cases will self-isolate or self-quarantine after 2 days but do not give any reference or justification for this assumption. They assume that the population of Ireland is 4, 900, 000 and state that 4, 977, 400 is a more accurate value from the Central Statistics Office but do not give any justification for not using the latter, the more accurate value.	
	4. Are the input parameters reasonable?	No to minor concerns	
	4. Comments	Input parameters seem to be appropriate as far as it can be assessed.	
Validation (external)	5. Has an external validation process been described?	Not reported	
	5. Comments	No external validation	
	6. Has the model been shown to be externally valid?	Moderate concerns	
	6. Comments	No external validation	
Validation (internal)	7. Has an internal validation process been described?	Not reported	
	7. Comments	No internal validation reported.	
	8. Has the model been shown to be internally valid?	Moderate concerns	
	8. Comments	No internal validation	
Uncertainty	9. Was there an adequate assessment of the effects of uncertainty?	Moderate concerns	
	9. Comments	Authors present a number of case studies in which some, but far from all, parameters are varied.	
Transparency	10. Was technical documentation, in sufficient detail to allow (potentially) for replication, made available openly or under agreements that protect intellectual property?	Major concerns	

(Continued)

		10. Comments	Although the authors describe the model equations in detail, the methods used to fit these model equations are not described in sufficient detail to be able to reproduce the results.
James 2020	Model structure	1. Are the structural assumptions transparent and justified?	No to minor concerns
		1. Comments	
		2. Are the structural assumptions reasonable given the overall objective, perspective and scope of the model?	No to minor concerns
		2. Comments	
	Input data	3. Are the input parameters transparent and justified?	Moderate concerns
		3. Comments	Not all parameters are justified: for instance, no justification for adding a normally distributed random variable with mean 0 and standard deviation 1 to the date recorded by case recall to obtain symptom onset date or for adding a Gamma distributed random variable with mean 6.7 and standard deviation 5.4 to the reported date for cases with no onset date reported.
		4. Are the input parameters reasonable?	No to minor concerns
		4. Comments	
	Validation (external)	5. Has an external validation process been described?	Not reported
		5. Comments	No external validation
		6. Has the model been shown to be externally valid?	Moderate concerns
		6. Comments	No external validation
	Validation (internal)	7. Has an internal validation process been described?	Not reported
		7. Comments	No internal validation reported.
		8. Has the model been shown to be internally valid?	Moderate concerns
		8. Comments	No internal validation
	Uncertainty	9. Was there an adequate assessment of the effects of uncertainty?	Major concerns

(Continued)

		9. Comments	No sensitivity analyses or other assessments of uncertainty.
	Transparency	10. Was technical documentation, in sufficient detail to allow (potentially) for replication, made available openly or under agreements that protect intellectual property?	Moderate concerns
		10. Comments	Code not available, methods not described in sufficient detail to allow others to replicate the analysis well.
Kang 2020	Model structure	1. Are the structural assumptions transparent and justified?	Moderate concerns
		1. Comments	Structural assumptions related to the choice of variables to construct the synthetic control are listed transparently, but are not well justified.
		2. Are the structural assumptions reasonable given the overall objective, perspective and scope of the model?	Major concerns
		2. Comments	Factors considered important for the spread of SARS-CoV-2, and used to construct the synthetic control, included demographic factors, economic factors, the healthcare environment and the number of Chinese visitors. These factors are quite superficial and may not capture relevant differences across countries in constructing a valid control - potentially important factors may not be covered, such as how governments otherwise reacted, e.g. how early testing or contact tracing strategies were implemented, if mask-wearing was already part of established prevention.
	Input data	3. Are the input parameters transparent and justified?	No to minor concerns
		3. Comments	Input parameters in this case are the observed data, and these are transparently reported.
		4. Are the input parameters reasonable?	Major concerns
		4. Comments	No consideration is given to different rates of case ascertainment across countries, which, especially early in the pandemic, likely varied quite widely across the countries assessed.
	Validation (external)	5. Has an external validation process been described?	Reported
		5. Comments	The synthetic control methodology uses observed data to construct a synthetic control; thus the overlay of the observed data and synthetic control data

(Continued)

			in the pre-intervention period represents a form of external validation.
		6. Has the model been shown to be externally valid?	No to minor concerns
		6. Comments	The various figures show that the counterfactual in the pre-intervention periods fit the observed data well.
	Validation (internal)	7. Has an internal validation process been described?	Not reported
		7. Comments	No internal validation reported.
		8. Has the model been shown to be internally valid?	Moderate concerns
		8. Comments	No internal validation
	Uncertainty	9. Was there an adequate assessment of the effects of uncertainty?	Major concerns
		9. Comments	No concrete assessment of uncertainty; additionally, point estimates are provided with no measure of precision. Given the low numbers of cases in the study period, it is likely that there was substantial imprecision which is not clear.
	Transparency	10. Was technical documentation, in sufficient detail to allow (potentially) for replication, made available openly or under agreements that protect intellectual property?	Major concerns
		10. Comments	Code not available, methods for implementing the synthetic control design only poorly described; not likely that one could replicate the analysis well.
Liebig 2020	Model structure	1. Are the structural assumptions transparent and justified?	Moderate concerns
		1. Comments	The model to estimate traveller volumes assuming no travel restrictions is only described as a seasonal autoregressive integrated moving average model determined via step-wise search over the model space. Moreover, it is unclear why the historical data on number of arrivals into Australia are Box-Cox transformed to give the data a normal shape.
		2. Are the structural assumptions reasonable given the overall objective, perspective and scope of the model?	Moderate concerns
		2. Comments	The assumption that a country's incidence rate equals the percentage of observed COVID-19 in-

![Cochrane Library - Trusted evidence. Informed decisions. Better health.]

Cochrane Database of Systematic Reviews

(Continued)

fections amongst travellers arriving into Australia from a given country is somewhat questionable, as individuals with COVID-19 symptoms may both be less likely to travel and more likely to be infected. It is somewhat inconsistent that in the model for the expected number of importations, a Poisson variable is drawn instead of using directly the rate parameter of this distribution as this model also uses the probability that an individual is infected and the probability that an individual is infectious during the flight rather than sampling from a Bernoulli or Binomial distribution with these probabilities.

Input data	3. Are the input parameters transparent and justified?	**Moderate concerns**
	3. Comments	In the importation model, the authors assume that arrivals spent an average of 15 days in the source country prior to arrival without further reference or justification. Uncertainty estimates are lacking for most parameters.
	4. Are the input parameters reasonable?	No to minor concerns
	4. Comments	Input parameters seem to be appropriate as far as it can be assessed
Validation (external)	5. Has an external validation process been described?	Not reported
	5. Comments	No external validation
	6. Has the model been shown to be externally valid?	Moderate concerns
	6. Comments	No external validation
Validation (internal)	7. Has an internal validation process been described?	Not reported
	7. Comments	No internal validation reported
	8. Has the model been shown to be internally valid?	Moderate concerns
	8. Comments	No internal validation
Uncertainty	9. Was there an adequate assessment of the effects of uncertainty?	Major concerns
	9. Comments	The authors did not explore alternative input parameter values and model assumptions.
Transparency	10. Was technical documentation, in sufficient detail to allow (potentially) for replication, made available openly or under	Major concerns

(Continued)

			agreements that protect intellectual property?
		10. Comments	Code is not reported. Only part of the data is available. Replication would be difficult.
Linka 2020a	Model structure	1. Are the structural assumptions transparent and justified?	No to minor concerns
		1. Comments	
		2. Are the structural assumptions reasonable given the overall objective, perspective and scope of the model?	Moderate concerns
		2. Comments	The authors assume that the number of infections are equal to the difference between the confirmed cases and the recovered cases and deaths from data from the ECDC but from my understanding this assumption is not very reasonable as the reporting of recovered cases is unreliable and incomplete.
	Input data	3. Are the input parameters transparent and justified?	No to minor concerns
		3. Comments	
		4. Are the input parameters reasonable?	Moderate concerns
		4. Comments	Based on data on the latent and infectious periods A 1⁄4 1=a 1⁄4 2.56 days and C 1⁄4 1=c 1⁄4 17.82 days from 30 Chinese provinces (Peirlinck et al. 2020). Yields estimates for R0 of 8.7 in Austria and 6.0 in Germany.
	Validation (external)	5. Has an external validation process been described?	Reported
		5. Comments	Limited approach to validate model projections against data used to build the model.
		6. Has the model been shown to be externally valid?	Moderate concerns
		6. Comments	Simulated data seems to be quite far from the observed data (see Figure-2).
	Validation (internal)	7. Has an internal validation process been described?	Not reported
		7. Comments	No internal validation
		8. Has the model been shown to be internally valid?	Moderate concerns

(Continued)

		8. Comments	No internal validation
	Uncertainty	9. Was there an adequate assessment of the effects of uncertainty?	Major concerns
		9. Comments	No sensitivity analyses or other assessments of uncertainty
	Transparency	10. Was technical documentation, in sufficient detail to allow (potentially) for replication, made available openly or under agreements that protect intellectual property?	Moderate concerns
		10. Comments	Code not available
Linka 2020b	Model structure	1. Are the structural assumptions transparent and justified?	No to minor concerns
		1. Comments	Model structure is well stated and seems to be reasonable
		2. Are the structural assumptions reasonable given the overall objective, perspective and scope of the model?	Moderate concerns
		2. Comments	The authors assume that the number of infections is equal to the "difference between today's and yesterday's reported cases" thereby ignoring underreporting and reporting delay.
	Input data	3. Are the input parameters transparent and justified?	No to minor concerns
		3. Comments	Within the scope of the analysis no concerns.
		4. Are the input parameters reasonable?	No to minor concerns
		4. Comments	Input parameters seem to be appropriate as far as it can be assessed.
	Validation (external)	5. Has an external validation process been described?	Reported
		5. Comments	Limited approach to validate model projections against data used to build the model.
		6. Has the model been shown to be externally valid?	Moderate concerns
		6. Comments	Only dependent validation. Figure 6 and 7 show a poor model fit to the data from Newfoundland and Labrador as 95% uncertainty intervals only cover a small percentage of the observed data points.

(Continued)

	Validation (internal)	7. Has an internal validation process been described?	Not reported
		7. Comments	No internal validation reported
		8. Has the model been shown to be internally valid?	Moderate concerns
		8. Comments	No internal validation
	Uncertainty	9. Was there an adequate assessment of the effects of uncertainty?	Major concerns
		9. Comments	No sensitivity analyses or other assessments of uncertainty.
	Transparency	10. Was technical documentation, in sufficient detail to allow (potentially) for replication, made available openly or under agreements that protect intellectual property?	Moderate concerns
		10. Comments	Code not available. As the analyses are based on an open-access modelled passenger flow matrix by Huang et al. (2010, Plos One) and a risk flow model proposed by Gilbert et al. (2020, The Lancet) it could in theory be possible to replicate the analyses. However, the model formulas are not explicit enough to understand exactly how to link these two sources. In particular, the authors perform simulations to understand how risk flow will change with travel restrictions and in some of the scenarios they assume that the transmissibility rate changes from 1 to 0.5 but it is unclear how this transmissibility rate intervenes in the model as the authors do not define a transmissibility rate in the method section and neither Huang et al. (2010, Plos One) nor Gilbert et al. (2020, The Lancet) define a transmissibility rate.
Mandal 2020	Model structure	1. Are the structural assumptions transparent and justified?	No to minor concerns
		1. Comments	
		2. Are the structural assumptions reasonable given the overall objective, perspective and scope of the model?	Major concerns
		2. Comments	The authors assume that all infections go through an asymptomatic stage and a symptomatic stage; they assume that the exposed are infected from the start (but with less severity as when showing symptoms).

(Continued)

			The model does somewhat mix the incubation period and the asymptomatic cases into one mixed stage
Input data	3. Are the input parameters transparent and justified?	Moderate concerns
	3. Comments	From my understanding, it is unclear how the authors determine the parameters of their SIR model concerning recovery and death and the parameter values are not given. The authors assume that asymptomatic cases are only 0.1 or 0.5 times as infectious as symptomatic cases but do not give any references for these assumptions.
	4. Are the input parameters reasonable?	Major concerns
	4. Comments	The authors assume that all symptomatic COVID-19 cases are identified and that zero, 50 or 90 percent of asymptomatic cases are identified. These values seem very optimistic (other authors, for instance Clifford 2020 or Gostic 2015 assume more realistic and justified parameter values concerning the sensitivity of screening measures).
Validation (external)	5. Has an external validation process been described?	Not reported
	5. Comments	No external validation
	6. Has the model been shown to be externally valid?	Moderate concerns
	6. Comments	No external validation
Validation (internal)	7. Has an internal validation process been described?	Not reported
	7. Comments	No internal validation
	8. Has the model been shown to be internally valid?	Moderate concerns
	8. Comments	No internal validation
Uncertainty	9. Was there an adequate assessment of the effects of uncertainty?	Moderate concerns
	9. Comments	Limited sensitivity analyses on a few parameters
Transparency	10. Was technical documentation, in sufficient detail to allow (potentially) for replication, made available openly or under agreements that protect intellectual property?	Moderate concerns

Cochrane
Library

Trusted evidence.
Informed decisions.
Better health.

Cochrane Database of Systematic Reviews

(Continued)

		10. Comments	Model formulas available but from my point of view, it is not very clear how the authors calibrated their SIR model. No code available.
McLure 2020	Model structure	1. Are the structural assumptions transparent and justified?	No to minor concerns
		1. Comments	
		2. Are the structural assumptions reasonable given the overall objective, perspective and scope of the model?	No to minor concerns
		2. Comments	
	Input data	3. Are the input parameters transparent and justified?	No to minor concerns
		3. Comments	
		4. Are the input parameters reasonable?	No to minor concerns
		4. Comments	
	Validation (external)	5. Has an external validation process been described?	Not reported
		5. Comments	No external validation
		6. Has the model been shown to be externally valid?	Moderate concerns
		6. Comments	No external validation
	Validation (internal)	7. Has an internal validation process been described?	Not reported
		7. Comments	No internal validation reported
		8. Has the model been shown to be internally valid?	Moderate concerns
		8. Comments	No internal validation
	Uncertainty	9. Was there an adequate assessment of the effects of uncertainty?	Major concerns
		9. Comments	No sensitivity analyses or other assessments of uncertainty.
	Transparency	10. Was technical documentation, in sufficient detail to allow (potentially) for replication, made available openly or under	No to minor concerns

(Continued)

		agreements that protect intellectual property?	
		10. Comments	Data and code are available.
Nakamura 2020	Model structure	1. Are the structural assumptions transparent and justified?	Moderate concerns
		1. Comments	The authors combine methods of existing publications but the model formulas are not sufficiently explicit to understand exactly how these existing publications are linked. In particular, in the results section, the authors describe the results of scenarios in which the transmissibility rate is reduced but it is not clear how this parameter (which is not defined in the methods section of either the present article or the articles referenced as sources for the methodology) intervenes in the model.
		2. Are the structural assumptions reasonable given the overall objective, perspective and scope of the model?	Moderate concerns
		2. Comments	The risk flow model proposed by Gilbert et al. (2020, The Lancet) accounts for A_{ia}, defined as the probability of travelling form i to a, conditioned on travelling internationally from i, while the authors of the present study define the quantity A_{od} as the probability of travelling from origin (o) to destination (d) conditioned on travelling internationally. It is unclear whether the omission of the conditioning of travelling from i is on purpose or not and it is questionable whether the normalization described on page 41 still holds if this conditioning is omitted.
	Input data	3. Are the input parameters transparent and justified?	No to minor concerns
		3. Comments	Within the scope of the analysis no concerns.
		4. Are the input parameters reasonable?	Moderate concerns
		4. Comments	The number of confirmed cases in different countries are used as input data in the model without accounting for or discussing different ascertainment rates in these countries.
	Validation (external)	5. Has an external validation process been described?	Not reported
		5. Comments	No external validation
		6. Has the model been shown to be externally valid?	Moderate concerns
		6. Comments	No external validation

(Continued)

	Validation (internal)	7. Has an internal validation process been described?	Not reported
		7. Comments	No internal validation reported
		8. Has the model been shown to be internally valid?	Moderate concerns
		8. Comments	No internal validation
	Uncertainty	9. Was there an adequate assessment of the effects of uncertainty?	Major concerns
		9. Comments	No sensitivity analyses or other assessments of uncertainty.
	Transparency	10. Was technical documentation, in sufficient detail to allow (potentially) for replication, made available openly or under agreements that protect intellectual property?	Moderate concerns
		10. Comments	Code not available. As the analyses are based on an open-access modelled passenger flow matrix by Huang et al. (2010, Plos One) and a risk flow model proposed by Gilbert et al. (2020, The Lancet), it could in theory be possible to replicate the analyses. However, the model formulas are not explicit enough to understand exactly how to link these two sources. In particular, the authors perform simulations to understand how risk flow will change with travel restrictions and in some of the scenarios they assume that the transmissibility rate changes from 1 to 0.5 but it is unclear how this transmissibility rate intervenes in the model as the authors do not define a transmissibility rate in the method section and neither Huang et al. (2010, Plos One) nor Gilbert et al. (2020, The Lancet) define a transmissibility rate.
Nowrasteh 2020	Model structure	1. Are the structural assumptions transparent and justified?	Moderate concerns
		1. Comments	Structural assumptions related to the choice of variables to construct the synthetic control are listed transparently, but are not well justified.
		2. Are the structural assumptions reasonable given the overall objective, perspective and scope of the model?	Moderate concerns
		2. Comments	Factors considered important for the spread of SARS-CoV-2, and used to construct the synthetic control, included immigrant population, the total population, population density, the percent of the population that is elderly, the share of the popu-

(Continued)

			lation that is urban, the median age, real GDP per capita (PPP), the absolute latitude or distance from the equator, the number of immigrants from China, and the number of airports with direct flights to China. This is a fairly comprehensive list, however some important aspects, such as how early testing or contact tracing strategies were implemented, if mask-wearing was already part of established prevention, may not have been picked up.
Input data	3. Are the input parameters transparent and justified?	No to minor concerns	
	3. Comments	Input parameters in this case are the observed data, and these are transparently reported.	
	4. Are the input parameters reasonable?	Major concerns	
	4. Comments	No consideration is given to different rates of case ascertainment across countries, which, especially early in the pandemic, likely varied quite widely across the countries assessed.	
Validation (external)	5. Has an external validation process been described?	Reported	
	5. Comments	The synthetic control methodology uses observed data to construct a synthetic control; thus the overlay of the observed data and synthetic control data in the pre-intervention period represents a form of external validation.	
	6. Has the model been shown to be externally valid?	No to minor concerns	
	6. Comments	The various figures show that the counterfactual in the pre-intervention periods fit the observed data well.	
Validation (internal)	7. Has an internal validation process been described?	Not reported	
	7. Comments	No internal validation reported.	
	8. Has the model been shown to be internally valid?	Moderate concerns	
	8. Comments	No internal validation	
Uncertainty	9. Was there an adequate assessment of the effects of uncertainty?	No to minor concerns	
	9. Comments	For each outcome four additional specifications were conducted to test the robustness of the results; these included changing the pre- and post-intervention periods.	

(Continued)

	Transparency	10. Was technical documentation, in sufficient detail to allow (potentially) for replication, made available openly or under agreements that protect intellectual property?	No to minor concerns
		10. Comments	Code available and methodology described to an extent that the study could be replicated.
Nuckchady 2020	Model structure	1. Are the structural assumptions transparent and justified?	No to minor concerns
		1. Comments	Model assumptions and equations are sufficiently stated in text and supplement; appendix C provides the model data.
		2. Are the structural assumptions reasonable given the overall objective, perspective and scope of the model?	Major concerns
		2. Comments	Simplistic model with reasonable assumptions on the topic at hand, but no adequate representation of reality (e.g. infected independent of severity of symptoms are equally infectious; homogenous risk of infection (no network effect); different population and age groups are not represented in spread and mortality data). Asymptomatic individuals do not get to the hospital and therefore do not get tested. Only 10% of symptomatic go to the hospital and have a chance of getting tested. No asymptomatic but infectious phase? Those without a test do not quarantine? All infected travelers who enter the country are assumed to be at the start of their incubation period. The model assumes that only infected individuals get tested, which has implications for the testing capacity.
	Input data	3. Are the input parameters transparent and justified?	No to minor concerns
		3. Comments	Parameters are transparently described based on the literature, real world data or reasoning based on real world data or literature. Some are assumptions, but this is made transparent.
		4. Are the input parameters reasonable?	Major concerns
		4. Comments	The test sensitivity is likely too low; given that primarily symptomatic individuals get tested. Assumptions regarding contact-behavior is likely overestimated - assumes no behavior change without intervention. With intervention, behavior change leading to a reduction of 20% is likely relatively low.

(Continued)

			Authors state: "These variables were manually modified to make the model fit the actual data." Indicating at an adjustment of model data post-hoc being in line with the real world figures.
	Validation (external)	5. Has an external validation process been described?	Reported
		5. Comments	Model was assessed against case figures from Mauritius.
		6. Has the model been shown to be externally valid?	Moderate concerns
		6. Comments	Model was assessed against case figures from Mauritius. While the real data is for the most part within the 95% confidence interval, the case estimates are no ideal fit. Author reports, that parameters were adjusted to fit the model: "These variables were manually modified to make the model fit the actual data" indicating issues with the external validation of the model
	Validation (internal)	7. Has an internal validation process been described?	Not reported
		7. Comments	No internal validation reported
		8. Has the model been shown to be internally valid?	Moderate concerns
		8. Comments	Process for internal validation not reported
	Uncertainty	9. Was there an adequate assessment of the effects of uncertainty?	Major concerns
		9. Comments	Author writes "multiple sensitivity analyses were conducted to ensure the results were robust." Unclear what this refers to; uncertainty for travel related outcome is not provided; insufficient variation of alternative scenarios / input parameters for the travel related measures.
	Transparency	10. Was technical documentation, in sufficient detail to allow (potentially) for replication, made available openly or under agreements that protect intellectual property?	Moderate concerns
		10. Comments	Code not reported, but inputs and assumptions reported sufficiently and should allow for replication.
Odendaal 2020	Model structure	1. Are the structural assumptions transparent and justified?	Moderate concerns

(Continued)

		1. Comments	Model intends to describe the impact of the travel restrictions in the US on 31 January 2020 (the interruption). It uses an exponential form to describe the cases outside China (as cases in China are considered unreliable).
		2. Are the structural assumptions reasonable given the overall objective, perspective and scope of the model?	Major concerns
		2. Comments	It is a very simple model that intends to describe the impact of the travel restrictions in the US on 31 January 2020 (the interruption). It uses a simple exponential form to describe the cases outside China (as cases in China are considered unreliable). It finds, that the number of cases (largely) follows an exponential growth rate (data fits better starting from mid-February 2020). The model actually does not predict the impact of the US travel restriction. The model does not conclude that the 26 day delay of community spread (the start of the exponential growth in the US) was caused by the US travel restrictions. They actually just postulate this: Observation 1: Community spread started around 26th of February 2020 (debatable); Starting here, the official number of cases in the US follow an exponential growth. Observation 2: the number of cases outside China largely follow an exponential growth rate Observation 3: the time between the implementation of the US travel restrictions (31 January 2020) and the beginning of community spread in the US is about 26 days. Conclusion: "The imposition of early travel restrictions from China into the USA slowed down the virus by containing it mainly to individuals who had been to infected areas. The model indicates that delay was 26 days before it reached "community-spread" in the USA. A main issue in this model is that it does not account for other confounders here. In particular: the cordon sanitaire around Wuhan on the 23 January 2020. It does not take travel patterns into account; it does not take the probability of seeding (regular travel patterns) into account. There is no causal inference; the 26 days are observed; not a results from the model
Input data	3. Are the input parameters transparent and justified?	No to minor concerns
		3. Comments	The parameters of the exponential growth model are derived from the official number of cases outside China. This is well justified and reported; the parameters for the exponential growth rates are given and justified

(Continued)

		The start of the community spread is based on official numbers in the US; the parameters for the exponential growth rate (starting around 29th) in the US are justified
	4. Are the input parameters reasonable?	No to minor concerns
	4. Comments	These parameters seem to be reasonable
Validation (external)	5. Has an external validation process been described?	Reported
	5. Comments	The model for the exponential growth of global cases is based on exported cases; model checked against these figures. No external validation specific to the outcome of interest "delay of community spread was conducted"
	6. Has the model been shown to be externally valid?	Moderate concerns
	6. Comments	The model for the exponential growth of global cases is based on exported cases. However, their model only starts to fit these figures starting from around mid-February 2020. The "global model" is not fitted against the US figures, rather, it reports that there is exponential growth starting around end of February 2020 as well.
Validation (internal)	7. Has an internal validation process been described?	Not reported
	7. Comments	No internal validation conducted
	8. Has the model been shown to be internally valid?	Moderate concerns
	8. Comments	No internal validation conducted
Uncertainty	9. Was there an adequate assessment of the effects of uncertainty?	Major concerns
	9. Comments	The study does not assess uncertainties and does not take other explanations into account
Transparency	10. Was technical documentation, in sufficient detail to allow (potentially) for replication, made available openly or under agreements that protect intellectual property?	Moderate concerns

(Continued)

		10. Comments	Code not available; the "model" and the input parameters are given and it should be possible to replicate the analysis
Pinotti 2020	Model structure	1. Are the structural assumptions transparent and justified?	No to minor concerns
		1. Comments	Model assumptions and equations are sufficiently stated in text and supplement
		2. Are the structural assumptions reasonable given the overall objective, perspective and scope of the model?	No to minor concerns
		2. Comments	
	Input data	3. Are the input parameters transparent and justified?	No to minor concerns
		3. Comments	There are no input parameters
		4. Are the input parameters reasonable?	No to minor concerns
		4. Comments	There are no input parameters
	Validation (external)	5. Has an external validation process been described?	Reported
		5. Comments	Dependent validation and predictive validation of the case arrival model
		6. Has the model been shown to be externally valid?	No to minor concerns
		6. Comments	Figure 3 shows good dependent validation and predictive performance of the case arrival model
	Validation (internal)	7. Has an internal validation process been described?	Not reported
		7. Comments	No internal validation reported
		8. Has the model been shown to be internally valid?	Moderate concerns
		8. Comments	No internal validation
	Uncertainty	9. Was there an adequate assessment of the effects of uncertainty?	Major concerns
		9. Comments	The authors did not explore alternative input parameter values and model assumptions
	Transparency	10. Was technical documentation, in sufficient detail to al-	Moderate concerns

(Continued)

			low (potentially) for replication, made available openly or under agreements that protect intellectual property?
		10. Comments	Model formulas are rather explicit but the code is not provided. Data is available
Quilty 2020	Model structure	1. Are the structural assumptions transparent and justified?	No to minor concerns
		1. Comments	The model is adequately described
		2. Are the structural assumptions reasonable given the overall objective, perspective and scope of the model?	Moderate concerns
		2. Comments	The model assumes 2 types of infections: asymptomatic and symptomatic of which 100% will experience a worsening of symptoms after 9.1 (+/- 14 days). It is unreasonable to assume that all patients will be hospitalized. It could however be assumed, that a share of those experiencing symptoms will decide not to travel; however it is unclear, why this should take place after 9.1 days, Furthermore, the study does not allow for self-reporting of symptoms but relies on the thermal scanners for detection.
	Input data	3. Are the input parameters transparent and justified?	No to minor concerns
		3. Comments	The study provides references almost all input parameters
		4. Are the input parameters reasonable?	Major concerns
		4. Comments	The study assumes that 83 % of infected will develop symptoms. In the light of the current evidence, this figure seems to be too high. The model assumes, that 100% of the patients will develop fever, which is not the case (there are infections which are symptomatic but do not have fever or only elevated temperature below the level to be detected by the thermal scanners. The sensitivity of the scanners was assumed to be 86%. In the study referenced, the value is given as 0.86% (95%CI 0.75–0.97) for a temperature of 37.8°. The assumption of 100% of symptomatic infected will refrain from travelling after mean 9.1 days (+/- 14.7 days) is likely too high.
	Validation (external)	5. Has an external validation process been described?	Not reported
		5. Comments	No external validation was conducted;

(Continued)

			The study does not discuss the findings in the light of empirical studies assessing screening interventions or the findings from repatriation studies
		6. Has the model been shown to be externally valid?	Moderate concerns
		6. Comments	No external validation was conducted; The study does not discuss the findings in the light of empirical studies assessing screening interventions or the findings from repatriation studies
	Validation (internal)	7. Has an internal validation process been described?	Not reported
		7. Comments	No internal validation conducted
		8. Has the model been shown to be internally valid?	Moderate concerns
		8. Comments	No internal validation conducted
	Uncertainty	9. Was there an adequate assessment of the effects of uncertainty?	Major concerns
		9. Comments	While the study itself does not provide a sensitivity analysis, it provides an App which should allow the reader to conduct sensitivity analyses; however the link provided for the app does not work
	Transparency	10. Was technical documentation, in sufficient detail to allow (potentially) for replication, made available openly or under agreements that protect intellectual property?	No to minor concerns
		10. Comments	Code to reproduce analysis is provided
Russell TW 2020	Model structure	1. Are the structural assumptions transparent and justified?	No to minor concerns
		1. Comments	Model structure is well stated and methods are found in another paper, also documented well. Model simplifications are discussed.
		2. Are the structural assumptions reasonable given the overall objective, perspective and scope of the model?	No to minor concerns
		2. Comments	Correction for under-ascertainment refines model results. Neglecting effects on local dynamics is reasonable if the relative number of imported cases is comparably low (which was found to be true for most countries).

(Continued)

			Treating all infected cases the same way does not account for symptom status of individuals. Other modes of travel will dominate in neighbouring countries. Both of these points were mentioned in the discussion.
Input data	3. Are the input parameters transparent and justified?	No to minor concerns	
	3. Comments	Source for case and death data is cited. Air travel data for slightly different scenarios is stated by the respective organization, but no direct source is given. Duration of infectiousness is stated as 10 days. Case Fatality ratio is cited, but not given.	
	4. Are the input parameters reasonable?	No to minor concerns	
	4. Comments	Duration of infectiousness not ideal when calculating the prevalence of infection (but probably no impact?); Reasonable otherwise.	
Validation (external)	5. Has an external validation process been described?	Reported	
	5. Comments	True case data is calibrated with available data (dependent validation), but calibration not illustrated.	
	6. Has the model been shown to be externally valid?	Moderate concerns	
	6. Comments	No actual validation besides calibration is available.	
Validation (internal)	7. Has an internal validation process been described?	Not reported	
	7. Comments	No internal validation reported	
	8. Has the model been shown to be internally valid?	Moderate concerns	
	8. Comments	Difficult part of code (under-ascertainment) has been published in previous paper and reports some external validation, implying some sense of validity; Rest of the model not too difficult to allow for many mistakes.	
Uncertainty	9. Was there an adequate assessment of the effects of uncertainty?	Moderate concerns	
	9. Comments	Main results were reported with credible intervals. Structural uncertainties were discussed, but not analysed quantitatively although they could arguably have an impact.	

			Origin of credible intervals not discussed, probably in methods paper? Sensitivity to flight data has been assessed by using different data sets. Sensitivity with respect to prevalence and incidence supposedly analysed, but not illustrated in paper.
	Transparency	10. Was technical documentation, in sufficient detail to allow (potentially) for replication, made available openly or under agreements that protect intellectual property?	No to minor concerns
		10. Comments	Data sufficiently cited to reconstruct their origin. Code is available.
Russell WA 2020	Model structure	1. Are the structural assumptions transparent and justified?	No to minor concerns
		1. Comments	Model assumptions and equations are sufficiently stated in text and supplement.
		2. Are the structural assumptions reasonable given the overall objective, perspective and scope of the model?	Moderate concerns
		2. Comments	The model only differentiates between infectious and non-infectious; not taking the peak-infectiousness (leading to super spreading events) in the time around symptom-development into account. The sensitivity of the test only distinguishes between symptomatic/presymptomatic state, rather than being reflective of the development of varying sensitivity during the course of the pandemic.
	Input data	3. Are the input parameters transparent and justified?	Moderate concerns
		3. Comments	Parameters are transparent; but limited justification and references for the selected parameters is provided.
		4. Are the input parameters reasonable?	Moderate concerns
		4. Comments	Due to the fixed sensitivity; the presymptomatic phase is relatively high (assuming, that days 1-3 of the infection PCR testing is very likely false negative). Sensitivity figures seem to refer to PCR testing (reference) although the figures in the cited publication do not match those reported here. If assumed for antigen-rather than PCR testing, this seems to be acceptable. The assumption of 40% non-infectiousness is likely too high (10.1371/journal.pmed.1003346).

(Continued)

			All figures can be changed in the online app.
	Validation (external)	5. Has an external validation process been described?	Not reported
		5. Comments	No external validation
		6. Has the model been shown to be externally valid?	Moderate concerns
		6. Comments	No external validation
	Validation (internal)	7. Has an internal validation process been described?	Not reported
		7. Comments	No internal validation reported
		8. Has the model been shown to be internally valid?	Moderate concerns
		8. Comments	Process for internal validation not reported
	Uncertainty	9. Was there an adequate assessment of the effects of uncertainty?	No to minor concerns
		9. Comments	Model introduced uncertainty by varying the distributions' mean and variance uniformly by ±20%, and sampled 1,000 parameter sets for the duration distributions. A number of alternative scenarios (high/low adherence; different levels of symptomatic status) are provided.
	Transparency	10. Was technical documentation, in sufficient detail to allow (potentially) for replication, made available openly or under agreements that protect intellectual property?	No to minor concerns
		10. Comments	Authors provide an app which allows to vary parameters. Calculations on parts of the model are provided as a supplement.
Ryu 2020	Model structure	1. Are the structural assumptions transparent and justified?	No to minor concerns
		1. Comments	It is a relatively simple SEIR model; it is not discussed why a more complex model was not used
		2. Are the structural assumptions reasonable given the overall objective, perspective and scope of the model?	Major concerns

(Continued)

	2. Comments	They assume all infected will become symptomatic; however this is not relevant for the model (as it is focused on an assumption of share of participants being quarantined); The model assumes, that those who will be quarantined, will be quarantined straight away, without having the ability to infect someone; The model does not allow for asymptomatic infectious; Model assumes random contacts within the whole population of Seoul; the interaction of the students is likely much more compartmentalized; Model assumes a perfect screening - no symptomatic infectious arrive	
Input data	3. Are the input parameters transparent and justified?	Moderate concerns	
	3. Comments	Model is based on the assumption of 0.1%, 0.2%, or 1% of students being infected. This is justified by: "(i) were in the pre-infectious period of COVID-19 infection, based on previous literature reporting that 0.2% of individuals with contactees of SARS infection were asymptomatic [11]" It is not clear, how the authors came up with these parameters	
	4. Are the input parameters reasonable?	Major concerns	
	4. Comments	The assumption of 70% - 100% of all infected being quarantined and not breaking quarantine seems very high, given that around 30-40% of all infections are asymptomatic The assumed incubation period seems a bit too long (6.5 days) The period from infectious to recovered is assumed to be 3.5 days. Within this model, the relevance of this parameters is the infectiousness period (only I are infectious). A period of 3.5 days infectiousness seems much too short.	
Validation (external)	5. Has an external validation process been described?	Reported	
	5. Comments	No external validation was conducted; The study does not discuss the findings in the light of empirical studies assessing screening interventions or the findings from repatriation studies	
	6. Has the model been shown to be externally valid?	Moderate concerns	
	6. Comments	No external validation was conducted; The study does not discuss the findings in the light of empirical studies assessing screening interventions or the findings from repatriation studies	

(Continued)

	Validation (internal)	7. Has an internal validation process been described?	Not reported
		7. Comments	No internal validation conducted
		8. Has the model been shown to be internally valid?	Moderate concerns
		8. Comments	No internal validation conducted
	Uncertainty	9. Was there an adequate assessment of the effects of uncertainty?	Moderate concerns
		9. Comments	The study varies the compliance rate wit quarantine (80%,80%,90%,100%) and the share of arriving infectious students (0.1%, 0.2%, 1%) but no further sensitivity analysis is conducted
	Transparency	10. Was technical documentation, in sufficient detail to allow (potentially) for replication, made available openly or under agreements that protect intellectual property?	Moderate concerns
		10. Comments	Code not available; the SEIR model and the input parameters are well described and it should be possible to replicate the analysis
Shi 2020	Model structure	1. Are the structural assumptions transparent and justified?	No to minor concerns
		1. Comments	Model assumptions are mostly established through existing literature. Method is based on another paper, but also summarized in this one. Choice of impact of travel restrictions with unclear formulation (75% of direct flights are cancelled). Missing explanation how other scenarios are evaluated, although this is probably just plugging in the estimated parameters.
		2. Are the structural assumptions reasonable given the overall objective, perspective and scope of the model?	Major concerns
		2. Comments	The concept of effective distance is non-trivial and could therefore impact results if not implemented correctly. Assumption that effective distance only from Wuhan is considered might be bad with other outbreak locations contributing to international spread (but assumption might be consistent with data time span until end of February 2020). Increasing number of cases in China increase the risk of exporting the virus but cannot be covered by

![Cochrane Library logo] Trusted evidence.
Informed decisions.
Better health.

Cochrane Database of Systematic Reviews

(Continued)

			effective distance. Therefore, time-constant hazard might be a poor assumption. Having many countries with increased risk of virus importation after imposing travel restrictions seems counter-intuitive and it is not sufficiently explained why the model produces these results.
Input data	3. Are the input parameters transparent and justified?	No to minor concerns
	3. Comments	Table with survival data and dates of travel restrictions is provided. Data source for airline network is stated. Additionally, data is available in a repository.
	4. Are the input parameters reasonable?	No to minor concerns
	4. Comments	There are no concerns regarding the input data
Validation (external)	5. Has an external validation process been described?	Reported
	5. Comments	Model is calibrated using data, therefore dependent validation is available. But calibration is barely illustrated.
	6. Has the model been shown to be externally valid?	Moderate concerns
	6. Comments	Although at least a dependent validation is available, there is no information about quality of estimates and model.
Validation (internal)	7. Has an internal validation process been described?	Not reported
	7. Comments	No internal validation is reported
	8. Has the model been shown to be internally valid?	Moderate concerns
	8. Comments	Model is simple in structure, not many concerns about internal validity. Similar results from the 25% and 50% travel reduction analysis provide some sense of validity.
Uncertainty	9. Was there an adequate assessment of the effects of uncertainty?	Major concerns
	9. Comments	No uncertainties for results are reported (providing quantiles is due to the nature of the results and not an analysis of uncertainty). There are many concerns whether structural assumptions are correct, analyses with alternative structures are necessary.

(Continued)

	Transparency	10. Was technical documentation, in sufficient detail to allow (potentially) for replication, made available openly or under agreements that protect intellectual property?	No to minor concerns
		10. Comments	Code and data are stored in accessible repositories.
Sruthi 2020	Model structure	1. Are the structural assumptions transparent and justified?	Moderate concerns
		1. Comments	Much of the structure is hidden away in an AI-type algorithm
		2. Are the structural assumptions reasonable given the overall objective, perspective and scope of the model?	Major concerns
		2. Comments	As far as it can be addressed the assumed structure seems reasonable; Many of the assumptions is impossible to assess given the information in the study
	Input data	3. Are the input parameters transparent and justified?	No to minor concerns
		3. Comments	Algorithm parameters are specified Not many more parameters as it seems
		4. Are the input parameters reasonable?	No to minor concerns
		4. Comments	Since model inputs are fairly straightforward, there are barely any problems A minor concern would be the input of recovery time which scales the reproduction rate
	Validation (external)	5. Has an external validation process been described?	Reported
		5. Comments	5-fold cross validation
		6. Has the model been shown to be externally valid?	Moderate concerns
		6. Comments	Cross-Validity seems to suggest that weekly infection rates can be predicted well if case numbers are high enough No other forms of validation reported
	Validation (internal)	7. Has an internal validation process been described?	Not reported
		7. Comments	No internal validation

(Continued)

		8. Has the model been shown to be internally valid?	Moderate concerns
		8. Comments	Functionality of cross-validation suggests that model is at least function in some sense
	Uncertainty	9. Was there an adequate assessment of the effects of uncertainty?	Moderate concerns
		9. Comments	Uncertainties were reported, but they likely do not span varying structural assumptions which may have significant impact on the reproduction rate contributions
	Transparency	10. Was technical documentation, in sufficient detail to allow (potentially) for replication, made available openly or under agreements that protect intellectual property?	No to minor concerns
		10. Comments	Code and source data available
Steyn 2020	Model structure	1. Are the structural assumptions transparent and justified?	No to minor concerns
		1. Comments	The model is described in adequate detail and had been used/described in a previous publication. The approach to simulate the passengers based on general epidemiological parameters is sensible in general
		2. Are the structural assumptions reasonable given the overall objective, perspective and scope of the model?	No to minor concerns
		2. Comments	Simplistic model with reasonable assumptions on the topic at hand.
	Input data	3. Are the input parameters transparent and justified?	No to minor concerns
		3. Comments	Parameters are transparently described based on the literature.
		4. Are the input parameters reasonable?	No to minor concerns
		4. Comments	All parameters (e.g. for rate of asymptomatic cases, incubation period, sensitivity of the PCR test in relation to time) are well described and justified with adequate literature.
	Validation (external)	5. Has an external validation process been described?	Not reported

(Continued)

		5. Comments	No external validation
		6. Has the model been shown to be externally valid?	Moderate concerns
		6. Comments	No external validation
	Validation (internal)	7. Has an internal validation process been described?	Not reported
		7. Comments	No internal validation reported
		8. Has the model been shown to be internally valid?	Moderate concerns
		8. Comments	Process for internal validation not reported
	Uncertainty	9. Was there an adequate assessment of the effects of uncertainty?	Major concerns
		9. Comments	The model does not explore alternative input parameter values and model assumptions
	Transparency	10. Was technical documentation, in sufficient detail to allow (potentially) for replication, made available openly or under agreements that protect intellectual property?	Moderate concerns
		10. Comments	Code not reported, but inputs and assumptions reported sufficiently and should allow for replication.
Taylor 2020	Model structure	1. Are the structural assumptions transparent and justified?	No to minor concerns
		1. Comments	Model is basic enough to fully develop the structure step by step; Most assumptions intuitive; Many small details make it difficult to assess
		2. Are the structural assumptions reasonable given the overall objective, perspective and scope of the model?	No to minor concerns
		2. Comments	Some minor concerns; Example: Probability of getting infected abroad equivalent to prevalence? Model might be too detailed for comprehensive analyses to be possible
	Input data	3. Are the input parameters transparent and justified?	No to minor concerns
		3. Comments	Parameter table is given including sources Assumed values are clearly stated

(Continued)

		4. Are the input parameters reasonable?	Moderate concerns
		4. Comments	Most parameters seem reasonable But high asymptomatic proportion Important parameters are neither reported with uncertainties nor backed up with enough sources (disease courses, non-compliance, efficiency of screening measures)
	Validation (external)	5. Has an external validation process been described?	Not reported
		5. Comments	No external validation
		6. Has the model been shown to be externally valid?	Moderate concerns
		6. Comments	No external validation
	Validation (internal)	7. Has an internal validation process been described?	Not reported
		7. Comments	No internal validation
		8. Has the model been shown to be internally valid?	Moderate concerns
		8. Comments	No internal validation
	Uncertainty	9. Was there an adequate assessment of the effects of uncertainty?	Major concerns
		9. Comments	Stochastic effects were considered Key uncertainties in disease characteristics, test efficiencies and structure of compliance were not investigated Existence of many parameter inputs requires a discussion of uncertainties
	Transparency	10. Was technical documentation, in sufficient detail to allow (potentially) for replication, made available openly or under agreements that protect intellectual property?	No to minor concerns
		10. Comments	Code not reported, but inputs and assumptions reported sufficient to allow for replication
Utsunomiya 2020	Model structure	1. Are the structural assumptions transparent and justified?	No to minor concerns
		1. Comments	Assume that the progression of COVID-19 for each country can be fit to a sigmoidal function, at either the lagging, exponential, decelerating or stationary stage; they also added a fifth more flexible stage

			to allow for bi-directional changes due to public health interventions, for example
	2. Are the structural assumptions reasonable given the overall objective, perspective and scope of the model?	Moderate concerns
	2. Comments	It is unclear whether a completely data-driven approach imposing a function onto the data is a reasonable approach; ignoring completely the transmission characteristics of COVID-19 as well as human behavior and mobility may be problematic.
Input data	3. Are the input parameters transparent and justified?	No to minor concerns
	3. Comments	No real input parameters required for this data-driven approach, apart from the data on the daily cases of COVID-19
	4. Are the input parameters reasonable?	No to minor concerns
	4. Comments	
Validation (external)	5. Has an external validation process been described?	Reported
	5. Comments	External validity assessed through predicting ECDC data 1-day in advance
	6. Has the model been shown to be externally valid?	No to minor concerns
	6. Comments	Figure 3 shows that the model was able to accurately predict the ECDC data
Validation (internal)	7. Has an internal validation process been described?	Reported
	7. Comments	Internal validation through the conduct of a simulation study; accuracy of estimates obtained by model was evaluated
	8. Has the model been shown to be internally valid?	No to minor concerns
	8. Comments	Figure-2 shows the results of the internal validity assessment
Uncertainty	9. Was there an adequate assessment of the effects of uncertainty?	Major concerns
	9. Comments	No real assessment of uncertainty through sensitivity analyses

Cochrane Library
Trusted evidence.
Informed decisions.
Better health.

Cochrane Database of Systematic Reviews

(Continued)

	Transparency	10. Was technical documentation, in sufficient detail to allow (potentially) for replication, made available openly or under agreements that protect intellectual property?	No to minor concerns
		10. Comments	Code and further information available at a linked Github repository
Kwok 2020	Model structure	1. Are the structural assumptions transparent and justified?	Moderate concerns
		1. Comments	Model structure was reported clearly and justified if necessary. Not exactly clear which local patches were modelled, by the information given seems like Hong Kong, Guangdong and China excluding Hubei are the patches considered. Temperature dependence of R0 needs more proof because it is quite controversial and has a high impact on results.
		2. Are the structural assumptions reasonable given the overall objective, perspective and scope of the model?	Major concerns
		2. Comments	Model seems to be simple to capture important effects. Assumption about R0 decreasing linearly from some initial value to zero due to temperature is highly questionable. Basically SEIR model in Hong Kong with varying R0 and possible influx from outside cases.
	Input data	3. Are the input parameters transparent and justified?	Moderate concerns
		3. Comments	Input parameters seem to be mentioned across the document, but no complete list. Parameters are stated without uncertainties. Rates for movement between patches unclear (is there outward flux from Hong Kong?).
		4. Are the input parameters reasonable?	Moderate concerns
		4. Comments	Population of Hong Kong is wrong by a factor 10, which might be a typo? Inputs for uncertainty of effectiveness of screening measures are necessary because of their high impact on results.
	Validation (external)	5. Has an external validation process been described?	Not reported
		5. Comments	No external validation

(Continued)

		6. Has the model been shown to be externally valid?	Moderate concerns
		6. Comments	No external validation
	Validation (internal)	7. Has an internal validation process been described?	Not reported
		7. Comments	No internal validation
		8. Has the model been shown to be internally valid?	Moderate concerns
		8. Comments	No internal validation
	Uncertainty	9. Was there an adequate assessment of the effects of uncertainty?	Major concerns
		9. Comments	Sensitivity in change of initial R0 was assessed. All other uncertainty analyses are missing (effectiveness of screening, change of dynamical model parameters). Since linear decrease of R0 is a critical but not sufficiently motivated assumption, modifications of the model structure should have been assessed due to the expected high impact of results.
	Transparency	10. Was technical documentation, in sufficient detail to allow (potentially) for replication, made available openly or under agreements that protect intellectual property?	Moderate concerns
		10. Comments	Code not reported and model parameters are in some cases unclear.
Wells 2020	Model structure	1. Are the structural assumptions transparent and justified?	No to minor concerns
		1. Comments	
		2. Are the structural assumptions reasonable given the overall objective, perspective and scope of the model?	No to minor concerns
		2. Comments	Assumed that no infected individuals travelled from Wuhan after the travel lockdown enforced on 23 January 2020
	Input data	3. Are the input parameters transparent and justified?	No to minor concerns
		3. Comments	

(Continued)

		4. Are the input parameters reasonable?	Moderate concerns
		4. Comments	Assumes that all symptomatic cases are identified in screening; Assumes a very high effectiveness for self-reporting of exposure risk; Based on early data on transmission parameters; assumes that the maximum incubation period is 21 days; that all reported infected cases acquired infection within mainland China
	Validation (external)	5. Has an external validation process been described?	Reported
		5. Comments	Predicted arrival times compared against observed
		6. Has the model been shown to be externally valid?	No to minor concerns
		6. Comments	Predicted first arrival times are generally consistent with reported international importation arrival dates
	Validation (internal)	7. Has an internal validation process been described?	Not reported
		7. Comments	No internal validation
		8. Has the model been shown to be internally valid?	Moderate concerns
		8. Comments	No internal validation
	Uncertainty	9. Was there an adequate assessment of the effects of uncertainty?	No to minor concerns
		9. Comments	
	Transparency	10. Was technical documentation, in sufficient detail to allow (potentially) for replication, made available openly or under agreements that protect intellectual property?	No to minor concerns
		10. Comments	Code available on GitHub and extensive supplementary material available.
Wilson 2020	Model structure	1. Are the structural assumptions transparent and justified?	No to minor concerns
		1. Comments	SEIR based model; with different assumptions and modelled interventions (e.g. wearing masks based on flight) for the parameter of influx of infected cases to protected, disease-free region (NZ)

(Continued)

		2. Are the structural assumptions reasonable given the overall objective, perspective and scope of the model?	Major concerns
		2. Comments	Model describes time until outbreak assuming that there is only one route of influx. The model does not assume any relevant countermeasures following a detected infection on a flight, which is unlikely. They assume that the only entry point to NZ would be Australia. And if: it is unlikely that the base risk is equal to the base risk of Australia
Input data		3. Are the input parameters transparent and justified?	No to minor concerns
		3. Comments	The study provides references for most assumptions
		4. Are the input parameters reasonable?	Major concerns
		4. Comments	While most assumptions on the variables are reported; there are a number of uncertainties in the underlying data. In particular, the assumptions for the effectiveness of the measures seem relatively arbitrary; e-g- on the effectiveness of entry and/or exit screening
Validation (external)		5. Has an external validation process been described?	Not reported
		5. Comments	No external validation was conducted; The study does not discuss the findings in the light of empirical studies assessing screening interventions or the findings from repatriation studies
		6. Has the model been shown to be externally valid?	Moderate concerns
		6. Comments	No external validation was conducted; The study does not discuss the findings in the light of empirical studies assessing screening interventions or the findings from repatriation studies
Validation (internal)		7. Has an internal validation process been described?	Not reported
		7. Comments	No internal validation conducted
		8. Has the model been shown to be internally valid?	Moderate concerns
		8. Comments	No internal validation conducted

(Continued)

	Uncertainty	9. Was there an adequate assessment of the effects of uncertainty?	Major concerns
		9. Comments	No to very limited sensitivity analyses
	Transparency	10. Was technical documentation, in sufficient detail to allow (potentially) for replication, made available openly or under agreements that protect intellectual property?	Moderate concerns
		10. Comments	Code not available; the SEIR model and the input parameters are well described and it should be possible to replicate the analysis
Wong MC 2020	Model structure	1. Are the structural assumptions transparent and justified?	Moderate concerns
		1. Comments	Authors describe their study methodology by relying heavily on citing another paper (Thompson et al. 2019); however, even after checking this publication it is not possible to understand completely what the structure of the model and analysis is.
		2. Are the structural assumptions reasonable given the overall objective, perspective and scope of the model?	Moderate concerns
		2. Comments	Given the poor description, it is not feasible to completely assess the methodology of this paper; assuming the authors remained close to the models described in Thompson et al. 2019, the structural assumptions are likely appropriate.
	Input data	3. Are the input parameters transparent and justified?	No to minor concerns
		3. Comments	Inputs for this estimation comprise the early observed cases in Hong Kong as well as an estimated serial interval; these are described sufficiently.
		4. Are the input parameters reasonable?	No to minor concerns
		4. Comments	No concerns related to these inputs.
	Validation (external)	5. Has an external validation process been described?	Reported
		5. Comments	Figure-2 shows the observed cases (with border control measures in place) versus the predicted cases (without measures in place), a comparison of the two curves in the period before control measures were in place serves as a form of dependent validation.

(Continued)

		6. Has the model been shown to be externally valid?	No to minor concerns
		6. Comments	The curves in Figure-2 are mostly consistent through the early stage of the pandemic, before measures were in place, suggesting the model is externally valid.
	Validation (internal)	7. Has an internal validation process been described?	Not reported
		7. Comments	No internal validation reported.
		8. Has the model been shown to be internally valid?	Moderate concerns
		8. Comments	No internal validation
	Uncertainty	9. Was there an adequate assessment of the effects of uncertainty?	Moderate concerns
		9. Comments	Authors write that the results are robust to the length of serial interval; however it does not appear that this is generalisable to the assessment of the impact of the control measures being assessed.
	Transparency	10. Was technical documentation, in sufficient detail to allow (potentially) for replication, made available openly or under agreements that protect intellectual property?	Major concerns
		10. Comments	Documentation of methods and code for the analysis is poor; authors reference another paper, however they provide very little detail on what they did, thus replicating this study would likely not be possible.
Yang 2020	Model structure	1. Are the structural assumptions transparent and justified?	No to minor concerns
		1. Comments	Model assumptions and equations are sufficiently stated in supplement. References for model analysis techniques are missing.
		2. Are the structural assumptions reasonable given the overall objective, perspective and scope of the model?	No to minor concerns
		2. Comments	Unclear if time-dependent parameters can be estimated reliably given the few observables. Travel restriction might be detrimental to country if there are strong internal infection dynamics; two-sided travel restrictions are assumed.

(Continued)

Input data	3. Are the input parameters transparent and justified?	No to minor concerns
	3. Comments	Parameters are reported with sources and additionally described in the main text Parameter tale in the supplement
	4. Are the input parameters reasonable?	No to minor concerns
	4. Comments	Description of parameters suggest that they are reasonable Critical mobility data well enough described Questionable inputs in would have no impact in many instances since they are only initials and are adjusted in fitting process
Validation (external)	5. Has an external validation process been described?	Reported
	5. Comments	Dependent validation on model predictions exists by construction
	6. Has the model been shown to be externally valid?	Moderate concerns
	6. Comments	Bare minimum of validation available by dependent validation
Validation (internal)	7. Has an internal validation process been described?	Not reported
	7. Comments	No internal validation reported
	8. Has the model been shown to be internally valid?	Moderate concerns
	8. Comments	Analyses on simulated data would have been great to chow that time-dependent parameter courses can indeed be estimated
Uncertainty	9. Was there an adequate assessment of the effects of uncertainty?	No to minor concerns
	9. Comments	Stochastic and parameter uncertainties are well covered by stochastic approach with adjustable parameter values Uncertainties on trajectories clearly visualized No analyses on model structure or mobility data
Transparency	10. Was technical documentation, in sufficient detail to allow (potentially) for replication, made available openly or under agreements that protect intellectual property?	No to minor concerns

(Continued)

		10. Comments	Python code and data are available Description of data analysis could have been more detailed
Zhang C 2020	Model structure	1. Are the structural assumptions transparent and justified?	No to minor concerns
		1. Comments	Linear Model is clearly stated and well explained. Implications of different outcome values are explained. There is some justification which time lag was assumed for different predictors. Fitting procedure has been described.
		2. Are the structural assumptions reasonable given the overall objective, perspective and scope of the model?	Moderate concerns
		2. Comments	Model is motivated well enough to be reasonable. Parameters in the linear model are a bit confusing, but interpretations are given. Suspicious that the daily new infections from one day ago are a non-significant predictor for the next day in too many cases. Results are by construction correlations, not clear to which extent causal relations can be extracted.
	Input data	3. Are the input parameters transparent and justified?	No to minor concerns
		3. Comments	Input data for flights, case data and country restrictions are stated.
		4. Are the input parameters reasonable?	No to minor concerns
		4. Comments	A minor issue would be that the analysis only accounts for confirmed cases. As discussed in the main text, data before the 22 January 2020 is missing for China. Incubation period of 14 days is quite long.
	Validation (external)	5. Has an external validation process been described?	Not reported
		5. Comments	No external validation
		6. Has the model been shown to be externally valid?	Moderate concerns
		6. Comments	No external validation
	Validation (internal)	7. Has an internal validation process been described?	Reported
		7. Comments	Replication of results by use of other flight data and another case data source.

(Continued)

		8. Has the model been shown to be internally valid?	No to minor concerns
		8. Comments	Since data is probably quite similar, this is a check of internal validity. Model seems to describe the data well, high R-Square (although some form of visualization would have been nice).
	Uncertainty	9. Was there an adequate assessment of the effects of uncertainty?	Moderate concerns
		9. Comments	Full table of all linear model results is given. P-values for parameter were reported, although not according to best practices (only inequalities, different thresholds). Using other predictors for the model (different time lags) would have enhanced the model credibility.
	Transparency	10. Was technical documentation, in sufficient detail to allow (potentially) for replication, made available openly or under agreements that protect intellectual property?	No to minor concerns
		10. Comments	Input sources have been cited and code for analysis is available in the supplement.
Zhang L 2020	Model structure	1. Are the structural assumptions transparent and justified?	Moderate concerns
		1. Comments	Model is minimalistic, the few equations used are defined. Variables are defined confusingly, difficult to exactly understand what they mean. Since connectivity is the central variable, its properties should have been explained more (adopted from other publication).
		2. Are the structural assumptions reasonable given the overall objective, perspective and scope of the model?	Major concerns
		2. Comments	Given the information, it was unclear why several things were done (estimation of cases on day n?, sum over the past 13 days when calculating imported cases on day n). Risk of a traveller being infected seems to be proportional to the cumulative cases of that country? (Crucial since this affects all results).
	Input data	3. Are the input parameters transparent and justified?	No to minor concerns

(Continued)

		3. Comments	Data sources have been cited, but it is seemingly a lot of data which should ideally be given in a supplement.
		4. Are the input parameters reasonable?	No to minor concerns
		4. Comments	Data sources seem to be appropriate as far as it can be assessed.
	Validation (external)	5. Has an external validation process been described?	Reported
		5. Comments	Figure-2 compares case risk index with imported cases.
		6. Has the model been shown to be externally valid?	Moderate concerns
		6. Comments	Unclear whether this is actual validation, since the imported cases seem to be estimated quite similarly as the case risk index. If the data of imported cases is actual data, this would be some form of validation, but this seems to be not the case.
	Validation (internal)	7. Has an internal validation process been described?	Not reported
		7. Comments	No form of internal validation was reported, but model is also quite simple.
		8. Has the model been shown to be internally valid?	Moderate concerns
		8. Comments	No large concerns because there is not much to validate.
	Uncertainty	9. Was there an adequate assessment of the effects of uncertainty?	Major concerns
		9. Comments	Uncertainty has not been considered.
	Transparency	10. Was technical documentation, in sufficient detail to allow (potentially) for replication, made available openly or under agreements that protect intellectual property?	Major concerns
		10. Comments	Code is not reported. Data must be aggregated from different sources. Replication would be difficult.
Zhong 2020	Model structure	1. Are the structural assumptions transparent and justified?	Moderate concerns

(Continued)

		1. Comments	Model structure is based on existing publication. Extensions are derived and explained in the supplement. Notation becomes complex but is summarized in a table. Arrival time and infected case reduction should have been defined more clearly. Sources are missing in supplement and are poorly cited in main document (Preprint version)
		2. Are the structural assumptions reasonable given the overall objective, perspective and scope of the model?	No to minor concerns
		2. Comments	Strong legitimation is given by a methods paper. Paper explains structure in detail, but reported analyses are a bit difficult to understand. Slope of linear relationship could change over time with more travel restrictions (was assumed constant?)
	Input data	3. Are the input parameters transparent and justified?	No to minor concerns
		3. Comments	No unreported parameters were noticed.
		4. Are the input parameters reasonable?	Moderate concerns
		4. Comments	There are some concerns with the nature of travel restriction parameters which have been assumed.
	Validation (external)	5. Has an external validation process been described?	Reported
		5. Comments	One small comparison of model prediction to independent value? Important model parameters were fitted to reproduce the linear relationship, some dependent validation.
		6. Has the model been shown to be externally valid?	Moderate concerns
		6. Comments	There was a comparison to real-world data, although I could not reconstruct the argument. Nevertheless, it would be only a weak validation.
	Validation (internal)	7. Has an internal validation process been described?	Reported
		7. Comments	The model was able to reproduce the important features on simulated data.
		8. Has the model been shown to be internally valid?	No to minor concerns

Cochrane
Library

Trusted evidence.
Informed decisions.
Better health.

Cochrane Database of Systematic Reviews

(Continued)

		8. Comments	Approach seems to technically work as intended.
	Uncertainty	9. Was there an adequate assessment of the effects of uncertainty?	Moderate concerns
		9. Comments	Important assumption of unchanging slope was analysed in sensitivity analysis. Main results are stated with uncertainties. Many smaller results reported without uncertainties. Travel restriction parameters should have been explored in sensitivity analyses. Appropriateness of model structure was partially discussed when results needed further explanation. There was a discussion of further possible uncertainties at the end.
	Transparency	10. Was technical documentation, in sufficient detail to allow (potentially) for replication, made available openly or under agreements that protect intellectual property?	Moderate concerns
		10. Comments	No code was reported and replication should be difficult, but seemingly possible.

Appendix 10. Travel restrictions reducing or stopping cross-border travel: study-by-study overview of the evidence contributing to each outcome (modelling studies)

Outcome	Number of studies	Overview of effect by study	Comparison used in each study	Effect direction per study (positive ▲; negative ▼; no change/mixed effects/conflicting findings ◄►)
Outcome category: cases avoided due to the measure				
Number or proportion of cases in the community	13 modelling studies	Anderson 2020: Across regions, relaxing border closures led to additional cases, the number of which varied between countries and world regions and based on the contact rate. The number of additional cases after six weeks seeded by one presymptomatic infectious traveller per week was higher for regions where substantial community transmission was occurring (e.g. California: 25 to 100 cases, Sweden: 20 to 80 cases), and lower for regions where less community transmission was occurring (e.g. Japan: 1 to 5 cases; New Zealand: 5 to 10 cases). Additionally, a lower contact rate - as a gener-	Maintaining versus relaxing measure	Positive (▲)

(Continued)

	al measure of the amount of social contact among a population - led to a lower number of additional cases in all regions (values estimated from the figure).		
	Banholzer 2020: Border closures across 12 high-income countries would have led to a 26% reduction in new cases (95% CI 13 to 37) compared with no border closure. According to the authors these results should be interpreted with caution due to the difficulty of disentangling the specific effects of various measures, and because of the critical role that the timing of the introduction of measures (which they did not assess) may play.	Measure versus no measure	Positive (▲)
	Binny 2020: In comparison to the real-life scenario, during which border restrictions were followed after 5 days by a border closure, which were followed by stricter community measures beginning 4 days later (1448 cumulative cases, 95% CI 1208 to 1796) and 4 daily cases, 95% CI 1 to 8), an early implementation of border restrictions or a delayed border closure would have led to 1422 (95% CI 1194 to 1765) and 1594 (95% CI 1359 to 1934) cumulative cases, and 4 (95% CI 1 to 8) and 5 (95% CI 1 to 9) daily cases, respectively. Implementing only border restrictions and border closures, yet not following up with stricter community-based measures such as a stay-at-home order and extensive testing and contact tracing would have led to sustained community transmission and an increase of 60443 cumulative cases (95% CI 45761 to 79201) and 1127 daily cases of (95% CI 841 to 1492).	Earlier versus later implementation of measure	Positive (▲)
	Chen T 2020: In China, the stringency of travel restrictions reducing the volume of travellers was seen to have an impact on the number of cases in community, with the strictest and most relaxed restrictions corresponding to 94 and 1148 cases, respectively. In Singapore, the stringency of travel restrictions, reducing the volume of travellers, also had a large impact on the number of cases in the community, with the strictest and most relaxed measures corresponding to 3042 and 44229 imported cases, respectively. Costantino 2020: A full ban on international travel from China followed by full relaxation was shown to lead to a reduction in the total number of cases in Australia (fewer than 300 cases compared to 2000 cases without the ban).	More versus less stringent measure	Positive (▲)
	Costantino 2020: A full ban on international travel from China followed by full relaxation was shown to lead to a reduction in the total number of cases in Australia (fewer than 300 cases compared to 2000 cases without the ban).	Measure versus no measure	Positive (▲)
	Deeb 2020: Limiting the number of international arrivals at the Beirut airport in Lebanon would have led to the arrival of 830 additional cases compared	More versus less stringent measure	Positive (▲)

(Continued)

to closing the airport completely. According to the authors, further relaxation of the airport closure would lead to a potentially large increase in the number of cases.

Kang 2020: Across all six countries assessed (Australia, Singapore, US, Vietnam, Taiwan, Hong Kong), the number of observed cumulative cases after the implementation of the travel ban on travellers from China was lower than the predicted number of cases if no travel ban had been introduced. The effects ranged from 81.3% reduction in the US (62 observed cases; 331 predicted cases) to 97.8% reduction in Vietnam (16 observed cases; 723 predicted cases).	Measure versus no measure	Positive (▲)
Linka 2020a: The introduction of travel restrictions across almost all European countries was shown to lead to a decrease in the proportion of infectious individuals in the population (no effect estimate available).	Measure versus no measure	Positive (▲)
Nowrasteh 2020: The study assessed the effect of banning the entry of all travellers who were physically present in China during the 14-day period preceding their entry or attempted entry into the US, with some exceptions for US permanent residents and those closely related to American citizens. The magnitude and direction of the difference between the cumulative and daily number of cases in the US (with travel restrictions) and the synthetic US (without travel restrictions) varied with the specification of the travel measure.	Measure versus no measure	Mixed (◀▶)
Kwok 2020: Without border closure, the number of cumulative cases in Hong Kong would vary with the level of community transmission, with a Rt = 2.2 associated with 29163 cases and a Rt = 1.6 associated with 2114 cases. A border closure between Hong Kong and China would lead to a 14% reduction in the number of cumulative cases in Hong Kong when community transmission was higher (Rt=2.2) and a 12% reduction in cases when community transmission was lower (Rt=1.6).	Measure versus no measure	Positive (▲)
Yang 2020: In general, across 13 high-income countries in Europe and North America, an earlier ban of international travel would lead to fewer daily cases. The magnitude, however, differs between countries; for some countries, e.g. Sweden and the UK, the reduction due to an earlier ban would be large; in others, e.g. Germany, the Netherlands, Denmark, Belgium and Italy, the reduction is smaller. Finally, in the US and Switzerland, the impact of the travel ban works in the opposite direction, with earlier implementation leading to higher case numbers, likely stemming from a higher prevalence in the community than in incoming travellers (no effect estimate available).	Earlier versus later implementation of measure	Mixed (◀▶)

(Continued)

		Zhang C 2020: The authors examine how the daily number of cases in the country implementing the travel restriction correlates with the cumulative cases in the restricted countries before versus after the restriction. Mixed effects across 22 countries in 6 continents are observed (no effect estimates available). For example, the restriction on travellers from China, did not influence the number of daily cases in the restricting countries. Similarly, restrictions on international travellers put in place by China, Iran, South Korea, Italy and Australia generally did not seem to influence the number of daily cases. As another example, however, restrictions put in place by the US against some countries (Germany, Portugal and South Africa) were effective, while restrictions against several other countries (Spain, Italy, France, Turkey, Brazil and Belgium) were not.	Measure versus no measure	Mixed (◄►)
		Zhong 2020: The study looked at travel-related control measures globally, estimating that, overall, these measures led to 5029 fewer cases. Most of this was due to decreased travel resulting from measures implemented in China, Hong Kong and Italy, with measures in other countries (Taiwan, Turkey, Spain, the US, Germany, Vietnam and Brazil) leading to smaller reductions. Lockdowns were shown to be much more effective in increasing the 'effective distance' between countries than partial or complete entry bans. The most effective measures were those that prevented passengers from exiting high community transmission regions or countries, such as Wuhan (China) and Italy (no effect estimate available).	Measure versus no measure	Positive (▲)
Number or proportion of imported or exported cases	9 modelling studies	Adekunle 2020: The Wuhan lockdown and restrictions on travel from China were shown to lead to 55 fewer COVID-19 cases imported from China into Australia compared with no such restrictions. The number of imported cases was reduced from 70 (expected) to 15 (observed) (79% reduction). Restrictions on travel from Iran, South Korea and Italy did not lead to a reduction in the number of cases imported from these countries to Australia.	Measure versus no measure	Mixed (◄►)
		Anzai 2020: Various restrictions of travel from China were shown to lead to 226 (95% CI 86 to 449) cases not being exported from China to other countries (70% reduction) compared with no restrictions.	Measure versus no measure	Positive (▲)
		Chen T 2020: In China, the stringency of travel restrictions reducing the volume of travellers was seen to have an impact on the number of imported cases, with the strictest and most relaxed restrictions corresponding to 369 and 28526 imported cases, respectively. In Singapore, the stringency of travel restrictions, reducing the volume of travellers, also influenced the number of imported cases, with the strictest and most relaxed measures	More versus less stringent measure	Positive (▲)

(Continued)

	corresponding to 64 and 3737 imported cases, respectively.		
	Chinazzi 2020: Compared to no restrictions, international travel restrictions on China would lead to an initial reduction in the number of cases imported from China globally, but by 1 March 2020 the number of cases had rebounded.	Measure versus no measure	Positive (▲)
	Costantino 2020: A travel ban on China followed by a full and partial relaxation led to 13 and 7 imported cases, respectively in Australia. No travel ban would have led to 122 imported cases between 26 January and 4 April 2020. Overall, the ban led to over 100 fewer cases imported from China to Australia, with most of the additional cases having been avoided in the first 6 weeks of the epidemic (80% reduction).	Measure versus no measure	Positive (▲)
	Liebig 2020: Travel bans imposed on foreigners travelling from the US, China, New Zealand, UK, Italy, France, Austria, Netherlands and Switzerland to Australia, with entry allowed for citizens and residents of Australia travelling from these countries, were shown to lead to reductions in the importation of cases into Australia between 18% (Switzerland) and 54% (Italy). Earlier implementation of the bans generally led to further reductions in the importation of cases, however the gain ranged from moderate to marginal.	Measure versus no measure; Earlier versus later implementation of measure	Positive (▲)
	McLure 2020: In a re-estimation of the data from the study by Costantino 2020 with changes in the model components, the travel ban on China followed by a full or partial relaxation was shown to lead to only 4 and 3 imported cases, respectively, in Australia. No travel ban would have led to 19 imported cases between 26 January and 4 April 2020.	Measure versus no measure	Positive (▲)
	Russell TW 2020: Relaxing international travel restrictions would have likely contributed meaningfully to local transmission in most countries (out of total 162 countries assessed). In May 2020, the proportion of countries where imports would have contributed to over 10% of cases ranged from 75% (95% CI 63 to 129) to 56% (95% CI 113 to 114), depending on whether a high (consistent with travel in May 2019) or low (consistent with reduced travel in May 2020) baseline travel scenario was considered. In September 2020, the proportion of countries in which imports would have contributed to over 10% of cases ranged from 35% (95% CI 22 to 112) to 23% (95% CI 8 to 85), depending on a high or low baseline travel scenario.	Maintaining versus relaxing measure	Positive (▲)
	Wells 2020: The lockdown of Wuhan and the rest of the Hubei Province would lead to 549 cases not being exported from China to other countries (81% reduction) compared with no lockdown.	Measure versus no measure	Positive (▲)

(Continued)

Number or proportion of deaths	3 modelling studies	Binny 2020: In comparison to the real-life scenario, during which border restrictions were followed after 5 days by a border closure, which were followed by stricter community measures beginning 4 days later (23 deaths, 95% CI 14 to 33), an early implementation of border restrictions or a delayed border closure would have led to 22 (95% CI 14 to 32) and 25 (95% CI 16, 35) deaths, respectively. Implementing only border restrictions and border closures, yet not following up with stricter community-based measures such as a stay-at-home order and extensive testing and contact tracing would have led to sustained community transmission and an increase of 1187 deaths (95% CI 891 to 1565).	Earlier versus later implementation of measure	Positive (▲)
		Costantino 2020: A full ban on international travel from China followed by full relaxation was shown to lead to a reduction in the total number of deaths in Australia (8 deaths compared to 400 deaths without the travel ban).	Measure versus no measure	Positive (▲)
		Kwok 2020: Without border closure, the number of deaths would vary with the level of community transmission, with a Rt = 2.2 associated with 400 deaths and a Rt = 1.6 associated with 35 deaths. A border closure between Hong Kong and China was shown to lead to 14% of reduction in deaths in Hong Kong when community transmission was higher (Rt=2.2) and 12% reduction in deaths when community transmission was lower (Rt=1.6).	Measure versus no measure	Positive (▲)
Risk of importation or exportation	3 modelling studies	Nakamura 2020: Compared with the status quo of no restrictions, restricting air travel between countries around the world would reduce the risk of importing and exporting infected persons. If community transmission of the virus is also reduced (e.g. through community measures) the risk could be further reduced. However, even with restrictions in place, there is still some risk of importation and exportation in countries that are heavily connected to the international travel network (e.g. China, the USA, Turkey, much of Europe) (no effect estimate was available).	Measure versus no measure	Positive (▲)
		Shi 2020: Compared with the assumed status quo of 75% restriction of flights from China to those countries in which restrictions were in place at the end of February 2020, scenarios of 50% and 25% restrictions and no restrictions at international airports (including with a focus on global traffic hubs) were shown to have mixed effects globally. Lessening restrictions led to an increased risk of importation at some airports, but a decreased risk at other airports. Taken together, the various scenarios show that basing travel restrictions on flight volume reductions creates a complex dynamic situation among the global air traffic network, and that reducing flight volumes can lead to reduced risk,	Measure versus no measure	Mixed (◀▶)

(Continued)

		but also to increased risk of case importation (no effect estimate available).		
		Zhang L 2020: A one-country, one-flight-per-week policy in China was shown to result in a lower imported risk index than the counterfactual scenario, in which the policy was not implemented (no effect estimate available).	Measure versus no measure	Positive (▲)

Outcome category: shift in epidemic development

Probability of eliminating the epidemic	1 modelling study	Binny 2020: In comparison to the real-life scenario, during which border restrictions were followed after 5 days by a border closure, which were followed by stricter community measures beginning 4 days later (0.66 probability of epidemic elimination), an early implementation of border restrictions or a delayed border closure would have had 0.66 and 0.55 probability of eliminating the epidemic, respectively. Implementing only border restrictions and border closures, yet not following up with stricter community-based measures, such as a stay-at-home order and extensive testing and contact tracing would have led to sustained community transmission and a reduction in probability of ending the epidemic from 0.66 to 0.	Earlier versus later implementation of measure	Mixed (◀▶)
Effective reproduction number	2 modelling studies	Linka 2020a: After the implementation of the travel restrictions in the EU, countries saw an inflection point (i.e. break point) in Rt; the duration of time until this inflection point varied, with a mean of 12.6 days.	Measure versus no measure	Positive (▲)
		Sruthi 2020: Over time the full closure of borders in Switzerland was shown to lead to a reduction in Rt by approximately 0.045. Partial relaxation, involving the opening of land borders, led to a subsequent increase in Rt of 0.177 (95% CI 0.175 to 0.178). Further relaxation, which allowed international travel followed by quarantine upon arrival, did not lead to a change in Rt.	Measure versus no measure; Maintaining versus relaxing measure	Mixed (◀▶)
Time to outbreak	6 modelling studies	Anzai 2020: Various restrictions of travel from China to other countries were found to lead to a delay in time of a major epidemic compared with no restrictions. With an R0 of 2.2 and 3.7, the delay was less than one day, with an R0 of 1.5, it was 1.25-2.4 days.	Measure versus no measure	Positive (▲)
		Davis 2020: Importations from China into USA likely only played a role in case importation very early in the pandemic; for states where the epidemic arrived later, few importations were from China, and the travel restrictions are likely partially responsible for this. However, cases were imported from a range of other countries, suggesting that the restrictions on China were insufficient. Implementing the travel restriction on China one week earlier	Earlier versus later implementation of measure	Positive (▲)

(Continued)

		would have led to a delay in community transmission by 2 days.		
		Grannell 2020: Compared to a situation in which the border is completely open and remains open throughout the pandemic, closing the land border between Northern Ireland and Ireland generally would have little effect on the time of the epidemic peak.	Measure versus no measure	Mixed (◀▶)
		Linka 2020b: Compared with no ban, banning travel from the US would delay the time until which 0.1% of the Canadian population was infected by between 58 and 85 days, depending on levels of community transmission (Rt = 1.35 and 1.16, respectively).	Measure versus no measure	Positive (▲)
		Odendaal 2020: The implementation of a ban on travel from China to the USA was shown to delay the community transmission in the USA by 26 days compared to no ban.	Measure versus no measure	Positive (▲)
		Zhong 2020: The study looked at travel-related control measures globally, estimating that, overall, these measures would lead to a delay in epidemic arrival of 16.69 days (95% CI 13.90 to 19.45). Most of this was due to the decreased travel resulting from the travel bans imposed on Wuhan (China) and Italy, with measures in other countries (US, Netherlands, Russia, Australia) leading only to small delays. Around half of the travel-related control measures implemented globally did not lead to a delay in epidemic arrival.	Measure versus no measure	Mixed (◀▶)
Risk of outbreak	2 modelling studies	Anzai 2020: Various restrictions of travel from China to other countries were found to lead to reductions in the risk of an outbreak, varying with R0 and the proportion of contacts traced. The largest reduction was 37% with an R0 of 1.5 and 50% of contacts traced. The smallest reduction was 1% with an R0 of 3.7 and 10% of contacts traced.	Measure versus no measure	Positive (▲)
		Boldog 2020: This study assessed the impact of no travel restrictions compared to varying degrees of travel restrictions in China. Thailand and Korea: restrictions could be effective in preventing an outbreak when the local R is low (e.g. 1.1); when the local R is higher (e.g. 2.2) or when the number of cases in China increases (e.g. 600,000 cumulative cases), a beneficial impact of restrictions becomes increasingly unlikely. US: even at lower numbers of cases in China (150,000), 25%, 50% and 75% travel restrictions were found to lead to a risk of a major outbreak of 80%, 65% and 45%, respectively. At higher numbers of cases in China (400.000), 25% and 50% restrictions had no impact, while 75% restrictions were associated with a risk of a major outbreak of approximately 85%.	Measure versus no measure; more versus less stringent measure	Mixed (◀▶)

		Canada: at lower numbers of cases in China, 25%, 50% and 75% travel restrictions yielded a risk of a major outbreak of 35%, 30% and 15%, respectively; at higher numbers of cases in China these risks were 80%, 70% and 45%, respectively.		
Number or proportion of cases at peak	2 modelling studies	Binny 2020: In comparison to the real-life scenario, during which border restrictions were followed after 5 days by a border closure, which were followed by stricter community measures beginning 4 days later (80 daily cases at peak, (95% CI 67 to 99)), an early implementation of border restrictions or a delayed border closure would lead to 79 (95% CI 67 to 97) and 91 (95% CI 77 to 100) daily cases at the epidemic peak, respectively. Implementing only border restrictions and border closures, yet not following up with stricter community-based measures, such as a stay-at-home order and extensive testing and contact tracing would have led to sustained community transmission and an increase of 47592 (95% CI 47240 to 47962) in daily cases at the peak.	Earlier versus later implementation of measure	Positive (▲)
		Grannell 2020: Compared to a situation in which the border is completely open and remains open throughout the pandemic, closing the land border between Northern Ireland and Ireland would generally lead to a lower proportion of individuals infected at the epidemic peak. The magnitude would differ (0.3% to 8%), however, depending on the level of community transmission, as influenced by the implementation and relaxation of public health measures, on each side of the border. This study also shows that if borders were open, the increases in cases in one country could lead to meaningful increases in cases in the other country.	Measure versus no measure	Positive (▲)
Epidemic growth acceleration	1 modeling study	Utsunomiya 2020: International travel controls would lead to a decrease in the growth acceleration of the epidemic progression across 62 countries (−6.05% change, P < 0.0001) compared with no travel controls.	Measure versus no measure	Positive (▲)
Exportation growth rate	1 modelling study	Pinotti 2020: Compared to no travel restrictions, both the lockdown of Hubei, representing a ban of all travel, as well as travel restrictions on China would to lead to a decrease in the growth rate of cases exported from Hubei and the rest of China, to the rest of the world.	Measure versus no measure	Positive (▲)

Outcome category: cases detected due to the measure

No contributing study

Appendix 11. Screening at borders: study-by-study overview of the evidence contributing to each outcome (modelling studies)

Outcome	Number of studies	Overview of effect by study	Comparison used in each study	Effect direction per study (positive ▲; negative ▼; no change/mixed effects/conflicting findings ◄►)
Symptom/exposure-based screening at borders				
Outcome category: cases avoided due to the measure				
Number or proportion of cases exported	1 modelling study	Wells 2020: Assuming that only 35.7% of symptomatic individuals are detected, the number of cases exported per day from China would reduce by 82% (95% CI 72 to 95) resulting from screening measures put in place across the world, compared with no screening measures.	Measure versus no measure	Positive (▲)
Outcome category: shift in epidemic development				
Time to outbreak	4 modelling studies	Clifford 2020a: Entry and exit screening, alone or combined, and measures to increase awareness and encourage appropriate responses would delay an outbreak in a hypothetical population. Assuming a sensitivity of 86%, if introduced at the beginning of an outbreak when very few infected individuals arrive, the measures would delay the outbreak by several days (ranging from 1 to 8 days). If introduced later, when more infected individuals arrive, the measures would do little to delay the outbreak (ranging from less than 1 to 1 day).	Measure versus no measure	Positive (▲)
		Mandal 2020: With Rt = 2.0, entry screening of symptomatic individuals would lead to a delay in reaching 1000 cases (2.7-day delay, from 45 to 47.7 days) in a hypothetical population compared to no screening. If screening could detect 50% and 90% of asymptomatics the delay would increase to 7.4 and 20 days, respectively. With higher community transmission (Rt = 4.0) these values of sensitivity are all lower.	Measure versus no measure	Positive (▲)
		Nuckchady 2020: Assuming one infected person entered Mauritius per day, entry or exit screening with a sensitivity of 64% would delay an outbreak by 9.7 days, and screening with a sensitivity of 100% by 20 days.	More versus less stringent measure	Positive (▲)
		Wilson 2020: Under the assumption of one flight per day (7.1% of normal travel volume) in a hypothetical disease-free area (modelled on New Zealand), exit screening alone with 50% sensitivity, would delay an outbreak by 0.5 years, from 1.7 years (95% CI 0.04 to 6.09) to 2.2 years (95% CI 0.6 to 8.11) compared with no screening.	Measure versus no measure	Positive (▲)

Cochrane
Library

Trusted evidence.
Informed decisions.
Better health.

Cochrane Database of Systematic Reviews

(Continued)

Risk of outbreak	1 modelling study	Nuckchady 2020: Assuming one infected person entered Mauritius per 100 days, entry screening with 100% sensitivity would reduce the probability of an outbreak within 3 months to 10% and screening with 50% sensitivity would reduce the probability to 48%.	More versus less stringent measure	Positive (▲)

Outcome category: cases detected due to the measure

Number or proportion of cases detected	4 modelling studies	Bays 2020: Entry screening of all arriving travellers would detect 0.8% of infected travellers in a hypothetical population in a limited exposure scenario (i.e. short-term stay in country of departure and short flight) and 12% of cases in a higher-exposure scenario (i.e. longer-term stay in country of departure and long flight). The effectiveness of entry screening would thus be influenced by the time window in which the exposure may have occurred (i.e. longer windows of exposure mean a higher likelihood that incubation may have occurred prior to departure) as well as the duration of the flight (i.e. longer flights increase the likelihood that symptoms develop during the flight and can thus be detected through screening).	Measure versus no measure	Positive (▲)
		Gostic 2020: With 25% of cases assumed to be subclinical, combined entry and exit screening using thermal scanners and self-reporting of exposure would detect 27% (95% CI 10 to 47) of cases in a hypothetical population. On their own, exit and entry screening would detect 17% (95% CI 3 to 33) and 20% (95% CI 7 to 40) of cases, respectively. As the proportion of subclinical cases increases the proportion of cases detected goes down; conversely, as the proportion of subclinical cases decreases, the proportion of cases detected goes up.	Measure versus no measure	Positive (▲)
		Quilty 2020: Assuming a sensitivity of 86% for thermal scanner-based screening and 17% of asymptomatic cases being undetectable, entry and exit screening combined and entry screening alone both detected 53% (95% CI 35 to 72) of cases in a hypothetical population; exit screening alone was comparatively less effective, detecting 44% (95% CI 33 to 56) of cases.	Measure versus no measure	Positive (▲)
		Taylor 2020: Entry screening of all incoming travellers in the UK would lead to the detection of 0.8% (95% CI 0.2 to 1.6) of cases when using thermal imaging scanners and 1.1% (955 CI: 0.4 to 2.1) of cases when using health checks. The proportion of cases detected would be lower when compared with the self-isolation of all incoming travellers (51.3% for 7 days of self-isolation; 78% for 14 days of self-isolation).	Measure versus no measure	Positive (▲)

Test-based screening at borders

Outcome category: cases avoided due to the measure

(Continued)

Proportion of secondary cases	1 modelling study	**Dickens 2020**: Compared with no measure targeting incoming travellers in a hypothetical population, testing all incoming travellers upon arrival, followed by the isolation of test-positives and requiring a negative test at the end of isolation would lead to a reduction in secondary cases of 88% (95% CI 87 to 89) for a 7-day isolation period and 92% (95% CI 92 to 93) for a 14-day isolation period.	Measure versus no measure	Positive (▲)
Proportion of imported cases	1 modelling study	**Dickens 2020**: Compared with no measure targeting incoming travellers in a hypothetical population, testing all incoming travellers upon arrival, followed by the isolation of test-positives and requiring a negative test at the end of isolation would lead to a reduction of 90% of imported cases for a 7-day isolation period and 92% for a 14-day isolation period. Testing all incoming travellers and refusing entry to test positives led to a reduction of 77%.	Measure versus no measure	Positive (▲)

Outcome category: shift in epidemic development

No contributing study

Outcome category: cases detected due to the measure

Days at risk of transmission	2 modelling studies	**Clifford 2020b**: Requiring a single PCR test upon arrival to the UK from EU countries would have led to 2.0 days at risk of transmission (95% CI 0 to 10.8). This is shorter than the days at risk of transmission for symptom/exposure-based entry screening alone (2.1 days at risk (95% CI 0 to 11.2)). Requiring an additional pre-flight test would slightly improve the effect of the PCR test upon arrival.	Measure versus alternative measure	Positive (▲)
		Russell WA 2020: Requiring all incoming travellers to test upon arrival in a hypothetical population would have led to 2.3 days at risk of transmission (95% CI 2.1 to 2.6). This is shorter than the days at risk of transmission for no measure at entry (2.6 days at risk, (95% CI 2.3 to 2.9)).	Measure versus no measure	Positive (▲)
Probability of releasing an infected individual into the community	2 modelling studies (**Clifford 2020b**, **Steyn 2020**)	**Clifford 2020b**: Requiring a single PCR test upon arrival to the UK from EU countries would reduce the risk of releasing an infected individual into the community compared with symptom/exposure-based entry screening alone (RR: 0.55, 95% CI 0.28 to 0.83). Requiring an additional pre-flight test would slightly improve the effectiveness of the PCR test upon arrival.	Measure versus alternative measure	Positive (▲)
		Steyn 2020: The probabilities of releasing an infected individual as a result of testing at departure and upon arrival in New Zealand were 48%, 50%, and 53% for scenarios assuming no, moderate, and high risk of transmission while travelling, respectively. These were higher compared with the probability of releasing an infected individual following a 14-day quarantine of all incoming travellers.	Measure versus no measure	Positive (▲)

Appendix 12. Study-by-study overview of the evidence contributing to each outcome for intervention categories 2 and 4 (observational studies) *(Continued)*

Study context	Screening approach	Approach to identifying cases	Study data
Al-Qahtani 2020			
All travellers arriving at Bahrain International Airport from highly endemic areas between 25 February and 14 March 2020	• Screening for symptoms (not further specified) and known contact with SARS-CoV-2 infected individual upon arrival	• PCR test upon arrival • Mandatory 14-day quarantine with daily monitoring for symptom development and PCR test of those who developed symptoms • PCR test at end of quarantine on day 14	No. of individuals being evaluated: 2714 No. of individuals screened positively: not reported No. of cases screened positively: 44 No. of cases identified in total: 188 Cases missed by screening: 144 Prevalence: 6.9% Proportion of cases detected: 23.4% Positive predictive value (PPV): Not calculated, as the total number of symptomatic individuals among non-cases was not reported.
	• PCR test upon arrival	• Mandatory 14-day quarantine with daily monitoring for symptom development and PCR test of those who developed symptoms • PCR test at end of quarantine on day 14	No. of individuals being evaluated: 2714 No. of individuals tested positively: NA[1] No. of cases tested positively: 136 No. of cases identified in total: 188 Cases missed by testing: 52 Prevalence: 6.9% Proportion of cases detected: 72.3% Positive predictive value (PPV): Not calculated, as those with a positive PCR test were considered true cases; no information was available to determine false positives.
	Combined measures • PCR test upon arrival • Mandatory 14-day quarantine with daily monitoring for symptom development and PCR test of those who developed symptoms	• PCR test at end of quarantine on day 14	No. of individuals being evaluated: 2714 No. of cases identified by PCR test upon arrival: 136 No. of additional cases identified during quarantine: 27 Cases identified in RT-PCR test prior to release from quarantine: 25 Prevalence: 6.9% Proportion of cases detected through combined measures: 86.7% Among cases with a negative PCR test upon arrival, proportion identified through symptom monitoring and PCR test during quarantine: 51.9%
Al-Tawfiq 2020			
Travellers returning to Saudi Arabia between 6 March and 7 June 2020,	• PCR test within 24 hours of arrival	• Repeated PCR test for those with symptoms during entry screening	No. of individuals being evaluated: 1928 No. of individuals tested positively: NA No. of cases positively: 14 No. of cases identified in total: 23

(Continued)

and quarantined in facilities operated by Johns Hopkins Aramco Healthcare (JHAH) which provides medical services for employees of the energy company Saudi Aramco and their dependents		• Mandatory 14-day quarantine with daily monitoring for symptom development and PCR test of those who developed symptoms • PCR test before the end of quarantine on day 12 or 13	Cases missed by testing: 9 Prevalence: 1.2% Proportion of cases detected: 60.9% Positive predictive value (PPV): Not calculated, as those with a positive PCR test were considered true cases; no data were available to determine false positives.

Arima 2020

Three evacuation flights from Hubei, China to Japan between 29 and 31 January 2020	• Symptom-focused clinical examination upon arrival (e.g. fever, respiratory illness)	• PCR test upon arrival • Repeated PCR test for those with symptoms during entry screening • Mandatory 14-day quarantine with daily monitoring for symptom development and PCR test of those who developed symptoms • PCR test at end of quarantine on day 14	No. of individuals being evaluated: 566 No. of individuals screened positively: 63 No. of cases screened positively: 4 No. of cases identified in total: 12* Cases missed by screening: 8 Prevalence: 2.1% Proportion of cases detected: 33.3% Positive predictive value (PPV): 6.2% *Results for PCR test at end of quarantine were pending for 14 individuals.
	• PCR test upon arrival	• Repeated PCR test for those with symptoms during entry screening • Mandatory 14-day quarantine with daily monitoring for symptom development and PCR test of those who developed symptoms • PCR test at end of quarantine on day 14	No. of individuals being evaluated: 566 No. of individuals tested positively: NA No. of cases screened positively: 7 No. of cases identified in total: 12* Cases missed by testing: 5 Prevalence: 2.1% Proportion of cases detected: 58.3% Positive predictive value (PPV): Not calculated, as those with a positive PCR test were considered true cases; no data were available to determine false positives. *Results for PCR test at end of quarantine were pending for 14 individuals.
	Combined measures • PCR test upon arrival • Mandatory 14-day quarantine with daily monitoring for symptom development and PCR test of those	• Repeated PCR testing for those identified as symptomatic in entry screening • PCR test at end of quarantine on day 14	No. of individuals being evaluated: 566 No. of cases identified by PCR test upon arrival: 7 No. of additional cases identified and retained during quarantine: 4 Cases identified in PCR test prior to release from quarantine: 1 Prevalence: 2.1% Proportion of cases detected through combined measures: 91.7%

(Continued)

	who developed symptoms		Among cases with a negative PCR test upon arrival, proportion identified through symptom monitoring and PCR test during quarantine: 80%

Chen J 2020

Single flight from Singapore to Hangzhou, China, on 24 January 2020, flagged as high risk due to several passengers having recently been in Wuhan, China	• Symptom screening upon arrival using a questionnaire-based interview on the presence of symptoms (including fever, dry cough, shivering, expectoration and headaches) and temperature measurement prior to deplaning)	• PCR test on day 2 • Mandatory 14-day (for those not originating in Wuhan) or 21-day (for those originating in Wuhan) quarantine with daily monitoring for symptom development and PCR test of those who developed symptoms • PCR test on day 13	No. of individuals being evaluated: 335 No. of individuals screened positively: not reported No. of cases screened positively: 3 No. of cases identified in total: 16 Cases missed by screening: 13 Prevalence: 4.8% Proportion of cases detected: 18.8% Positive predictive value (PPV): Not calculated, as the number of symptomatic individuals among non-cases is not reported.
	Combined measures • PCR test on day 2 • Mandatory 14-day (for those not originating in Wuhan) or 21-day (for those originating in Wuhan) quarantine with daily monitoring for symptom development and PCR test of those who developed symptoms	• Mandatory 14-day (for those not originating in Wuhan) or 21-day (for those originating in Wuhan) quarantine with daily monitoring for symptom development and PCR test of those who developed symptoms • PCR test on day 13	No. of individuals being evaluated: 335 No. of individuals tested positively: NA No. of cases tested positively: 11 No. of cases identified in total: 16 Cases missed by screening: 5 Prevalence: 4.8% Proportion of cases detected: 68.8% Positive predictive value (PPV): Not calculated, as those with a positive PCR test were considered true cases; no data were available to determine false positives.

Hoehl 2020

Single evacuation flight from Hubei, China to Frankfurt, Germany on 1 February 2020	• Screening before departure based on (i) symptoms of COVID-19, (ii) being a close contact of an infected person, or (iii) accompanying a person belonging to (i) or (ii) • Symptom screening upon arrival	• PCR test likely upon arrival (although timing is not reported clearly) • Mandatory 14-day quarantine with observation of symptom development and PCR test of those who developed symptoms	No. of individuals being evaluated: 126 No. of individuals screened positively: 11 (10+1)* No. of cases screened positively: 0 No. of cases identified in total: 2** Cases missed by screening: 2 Prevalence: 1.6% Proportion of cases detected: 0% Positive predictive value (PPV): 0% *10 individuals were identified prior to departure (two based on contact, 6 based on symptoms, 2 based on being an ac-

(Continued)

		companying person); one individual was identified upon arrival. **One individual declined to receive the PCR test.	

Kim 2020

Single evacuation flight from Tehran, Iran to Korea, via Dubai, on 19 March 2020	• Symptom screening using health status questionnaire and temperature measurement prior to flight transfer in Dubai • Symptom screening using health status questionnaire and temperature measurement upon arrival	• PCR test upon arrival • Mandatory 14-day quarantine with observation of symptom development and likely PCR test of those who developed symptoms (although none developed symptoms).	No. of individuals being evaluated: 80 No. of individuals screened positively: 2 No. of cases screened positively: 1 No. of cases identified in total: 1 Cases missed by screening: 0 Prevalence: 1.3% Proportion of cases detected: 100% Positive predictive value (PPV): 50%

Lagier 2020

Three evacuation flights from Wuhan, China to France, between 30 January and 14 February 2020	• Symptom screening during flight • PCR test upon arrival, 3-8 days after beginning of pre-departure quarantine in China	• Mandatory 14-day quarantine with monitoring for symptom development and PCR test of those who developed symptoms • PCR test on day 5, 7-12 days after beginning of pre-departure quarantine in China	No. of individuals being evaluated: 337 No. of individuals tested positively: NA No. of cases tested positively: 0 No. of cases identified in total: 0 Cases missed by screening: 0 Prevalence: 0% Proportion of cases detected: Not calculated, as no cases were identified Positive predictive value (PPV): Not calculated, as no cases were identified

Lio 2020

Single evacuation flight from Wuhan, China to Macao, China on 7 March 2020	• PCR test on day 2	• Mandatory 14-day quarantine with monitoring for symptom development • PCR test on day 7 • PCR test on day 13 • Assessment of Sera antibodies on day 14	No. of individuals being evaluated: 57 No. of individuals tested positively: NA No. of cases tested positively: 0 No. of cases identified in total: 0 Cases missed by screening: 0 Prevalence: 0% Proportion of cases detected: Not calculated, as no cases were identified Positive predictive value (PPV): Not calculated, as no cases were identified

Lytras 2020

Seven evacuation flights from Lon-	• Symptom screening based	• PCR test upon arrival	No. of individuals being evaluated: 783 No. of individuals screened positively: 1

(Continued)

| don, UK, Madrid and Barcelona, Spain, and Istanbul Turkey to Athens, Greece, between 20 March and 25 March 2020 | on the reporting of 'general or respiratory symptoms' upon arrival | | No. of cases screened positively: 1
No. of cases identified in total: 40*
Cases missed by screening: 39*

Prevalence: 5.1%*

Proportion of cases detected: 2.5%

Positive predictive value (PPV): 100%

*In total, 40 individuals had a positive PCR test upon arrival; among these, one was symptomatic upon arrival and four developed symptoms during the observational period. An additional 36 individuals with a negative initial PCR test developed symptoms during quarantine, but did not receive an additional PCR test. Some of these could be additional COVID-19 cases. |

Ng 2020

| Single evacuation flight from Wuhan, China to Singapore on 30 January 2020 | • Symptom screening based on temperature measurement upon departure*
• Symptom screening based on temperature measurement upon arrival | • PCR test of symptomatic individuals upon arrival
• Mandatory 14-day quarantine with daily monitoring for symptom development and PCR test of those who developed symptoms
• PCR test of most asymptomatic individuals on day 3
• PCR test of all individuals at the government quarantine facility on day 6 | No. of individuals being evaluated: (97) 94*
No. of individuals screened positively: 2*
No. of cases screened positively: 2
No. of cases identified in total: 4**
Cases missed by screening: 2

Prevalence: 4.3%

Proportion of cases detected: 50%

Positive predictive value (PPV): 100%

*Of 97 individuals, prior to departure three were found to be febrile and denied boarding. It is unclear if they subsequently received a diagnostic test and, if so, what the outcome of this test was. Two additional individuals were found symptomatic during the screening upon arrival.

** One of the individuals was reported as having an indeterminate result, but was nevertheless transferred to a hospital for quarantine; here this individual is considered a probable case. |
| | • Symptom screening based on temperature measurement upon departure*
• Symptom screening based on temperature measurement upon arrival
• PCR test of symptomatic individuals upon arrival
• PCR test of asymptomatic individuals on day 3 | • Mandatory 14-day quarantine with daily monitoring for symptom development and PCR test of those who developed symptoms
• PCR test of all individuals at the government quarantine facility on day 6 | No. of individuals being evaluated: (97) 94*
No. of cases tested positively: 4**, ***
No. of cases identified in total: 4**
Cases missed by testing: 0

Prevalence: 4.3%

Proportion of cases detected: 100%

Positive predictive value (PPV): Not calculated, as those with a positive PCR test were considered true cases; no data were available to determine false positives.

*Of 97 individuals, prior to departure three were found to be febrile and denied boarding. It is unclear if they subsequently received a diagnostic test and, if so, what the outcome of this test was. Two additional individuals were found symptomatic during the screening upon arrival. |

(Continued)

			** One of the individuals was reported as having an indeterminate result, but was nevertheless transferred to a hospital for quarantine; here this individual is considered a probable case. ***In the quarantine period between arrival and first PCR test, 6 individuals reported symptoms but were tested negative for SARS-CoV-2

Shaikh Abdul Karim 2020

Five evacuation flights from China, Iran, Italy and Indonesia to Malaysia between February and April 2020	• Visual triaging prior to departure • Symptom screening upon arrival ('health screening', not further specified) • PCR test upon arrival	• Mandatory 14-day quarantine without monitoring for symptom development • PCR test on day 13	No. of individuals being evaluated: 432 No. of individuals tested positively: NA No. of cases tested positively: 74 No. of cases identified in total: 82 Cases missed by screening: 8 Prevalence: 19.0%* Proportion of cases detected: 90.3% Positive predictive value (PPV): Not calculated, as those with a positive PCR test were considered true cases; no data were available to determine false positives. *Among the four evacuation flights from Wuhan, Iran and Italy the prevalence was 0.6% (2 cases); among the repatriates from Indonesia the prevalence was 64,5% (80/124)
	Combined measures • PCR test upon arrival • Mandatory 14-day quarantine without monitoring for symptom development	• PCR test on day 13	No. of individuals being evaluated: 432 No. of cases identified by PCR test upon arrival: 74 No. of additional cases identified and retained during quarantine: 0 Cases identified in PCR test prior to release from quarantine: 8 Prevalence: 19.0%* Proportion of cases detected through combined measures: 90.3% Among cases with a negative PCR test upon arrival, proportion identified through symptom observation during quarantine: 0% *Among the four evacuation flights from Wuhan, Iran and Italy the prevalence was 0.6% (2 cases); among the repatriates from Indonesia the prevalence was 64,5% (80/124)

Wong J 2020

All travellers arriving in Brunei between 21 March and 24 April 2020	• Symptom screening upon arrival (not further specified)	• PCR test upon arrival • Mandatory 14-day quarantine	No. of individuals being evaluated: 1396 No. of individuals screened positively: 16 No. of cases screened positively: 16 No. of cases identified in total: 30 Cases missed by screening: 14 Prevalence: 2.2% Proportion of cases detected: 53.3% Positive predictive value (PPV): Not calculated, as the number of symptomatic individuals among non-cases is not reported.

Cochrane
Library

Trusted evidence.
Informed decisions.
Better health.

Cochrane Database of Systematic Reviews

(Continued)

Yamahata 2020

Travellers aboard the Diamond Princess cruise ship in the port of Yokohama, Japan on 3 February 2020	• Screening based on (i) symptoms of COVID-19 and (ii) being a close contact of someone with symptoms	• PCR test of all individuals identified through screening • PCR test of all individuals on board with PCR over a three-week period after arrival • Mandatory 14-day quarantine and symptom observation after negative PCR test	No. of individuals being evaluated: 1396 No. of individuals screened positively: 31 No. of cases screened positively: 10 No. of cases identified in total: 696 Cases missed by screening: 686 Prevalence: 18.8% Proportion of cases detected: 1.4% Positive predictive value (PPV): 32.3%

Appendix 13. Quarantine: study-by-study overview of the evidence contributing to each outcome (modelling studies)

Outcome	Number of studies	Overview of effect by study	Comparison used in each study	Effect direction per study (positive ▲; negative ▼; no change/mixed effects/conflicting findings ◀▶)
Outcome category: cases avoided due to the measure				
Number or proportion of cases in the community	3 modelling studies	Chen Y-H 2020: With test-and-isolation, contact tracing, and general public mask-wearing and other social measures in place, strict quarantine of travellers (1 daily infection imported) would ensure that the number of daily domestic infections remains low in Taiwan (349 cases) over 90 days; without quarantine (10 daily infections imported), the number of daily domestic infections would increase steadily (3483 cases) over the same time period.	Measure versus no measure	Positive (▲)
		Chen T 2020: Quarantining travellers entering China would lead to fewer cases in the country. Specifically, quarantining all inbound travellers would lead to 79 local cases, while quarantining only symptomatic individuals would lead to 1317 local cases. No quarantine of travellers would lead to 1534 local cases. Similarly, quarantining travellers entering Singapore would lead to fewer cases in the country. Specifically, quarantining all inbound travellers would lead to 2272 local cases, while quarantining only symptomatic individuals would lead to 27934 local cases. No quarantine of travellers would lead to 66300 local cases.	Measure versus no measure	Positive (▲)

(Continued)

		Wong MC 2020: During the first wave of infections in Hong Kong, quarantine of travellers would likely lead to 450 fewer cumulative cases in the community compared with no quarantine, however there is substantial uncertainty surrounding this estimate. During the second wave, the same measure would lead to 1650 fewer cumulative cases in the community.	Measure versus no measure	Positive (▲)
Proportion of imported cases	1 modelling study	Dickens 2020: Compared with no measure targeting incoming travellers, quarantining all incoming travellers at a hypothetical point of entry would reduce the proportion of imported cases by 55% for a 7-day quarantine period and by 91% for a 14-day quarantine period.	Measure versus no measure	Positive (▲)
Number or proportion of cases seeded by imported cases	3 modelling studies	Dickens 2020: Compared with no measure targeting incoming travellers, quarantining all incoming travellers at a hypothetical point of entry would reduce the proportion of secondary cases by 30% (95% CI 24 to 41) for a 7-day quarantine period and 84% (95% CI 78 to 89) for a 14-day quarantine period.	Measure versus no measure	Positive (▲)
		James 2020: Compared with the time period during which no measures were in place (expected number of secondary cases caused by imports: 0.63 cases among adults (95% CI 0.48 to 0.84) to 0.87 cases among elderly (95% CI 0.6 to 1.23)), a 14-day self-isolation of all international arrivals to New Zealand would lead to reductions in the number of secondary infections caused by imported cases (0.46 cases among adults (95% CI 0.40 to 0.51); 0.28 cases among elderly (95% CI 0.18 to 0.42)). The number of cases seeded would be even lower when a 14-day government-mandated quarantine of all international arrivals was required (0 cases among adults (95% CI 0 to 0.5); 0 cases among elderly (95% CI 0 to 0.33)).	Measure versus no measure	Positive (▲)
		Ryu 2020: The number of cases seeded by quarantined students arriving in South Korea from China in late March 2020 would be lower with higher compliance to quarantine; the number of cases seeded, for high and low compliance, are 19 and 45, 40 and 72, and 184 and 277, with the arrival of 0.1%, 0.2%, and 1% of pre-infectious individuals, respectively.	More versus less stringent measure	Positive (▲)
Probability of an imported case not infecting anyone	1 modelling study	James 2020: Compared with the time period during which no measures were in place (adults: 0.69 probability (95% CI 0.62 to 0.75), elderly: 0.64 probability (95% CI 0.47 to 0.78)), a 14-day self-isolation of all international arrivals in New Zealand would lead to an increase in the probability that an imported case would not infect anyone (adults: 0.73 (95% CI 0.7 to 076); elderly: 0.78 (95% CI 0.69 to 0.87). This probability would be higher when a 14-day government-mandated quarantine is required	Measure versus no measure	Positive (▲)

(Continued)

| | | (adults: 1 (95% CI 0.61 to 1); elderly: 1 (95% CI 0.72 to 1)). | | |

Outcome category: shift in epidemic development

| Time to outbreak | 1 modelling study | Kivuti-Bito 2020: With a quarantine of all travellers entering Kenya, the peak of approximately 13 million active cases and 34,000 deaths, would occur after approximately 180 days for the base case of 75% quarantine effectiveness. Increasing the effectiveness of quarantine to 80% and 90% would delay the peak in active cases and deaths by 3.5 and 5.5 days, respectively. | More versus less stringent measure | Positive (▲) |

Outcome category: cases detected due to the measure

Days at risk of transmission	2 modelling studies	Clifford 2020b: Requiring a mandatory quarantine in the UK for incoming travellers from the EU would lead to different days that the travellers are at risk of transmitting the infection into the community depending on the length of quarantine (3-day quarantine: 2 days (95% CI 0 to 10.2); 14-day quarantine: 0 days (95% CI 0 to 0)). These are lower than the days at risk of transmission for symptom/exposure-based entry screening alone (2.1 days at risk (95% CI 0 to 11.2) Requiring a pre-flight test only would slightly improve the effectiveness of the quarantine.	Measure versus no measure	Positive (▲)
		Russell WA 2020: Requiring all incoming travellers to enter a quarantine at a hypothetical point of entry would lead to different days at risk of transmission for travellers depending on the length of quarantine (2-day quarantine: 1.8 days (95% CI 1.6 to 2.2); 14-day quarantine: 0.53 days (95% CI 0.46 to 0.6)). These are lower than the days at risk of transmission when no measure is at place (2.6 days at risk (95% CI 2.3 to 2.9)).	Measure versus no measure	Positive (▲)
Proportion of cases detected	1 modelling studies	Taylor 2020: Requiring travellers to self-isolate upon arrival in the UK would lead to detecting different proportion of cases, with the magnitude increasing with the number of days in isolation (7 days: 51% (95% CI 47 to 56) of cases detected; 14 days: 78% (95% CI 74 to 82) of cases detected). These are higher than the proportion of cases detected based through screening alone (with either thermal imaging scanners or health checks detecting 0.78% and 1.13% of cases, respectively).	Measure versus no measure	Positive (▲)
Probability of releasing an infected individual into community	3 modelling studies	Ashcroft 2020: A 10-day quarantine at a hypothetical point of entry is sufficient to prevent almost all further transmission from international travellers, regardless of the duration of the travel. For shorter quarantine periods, the exact impact depends on the duration of travel: for longer-duration travel (e.g. 10-14 days), a 5-day quarantine will prevent over 80% of further transmission (probability of releasing an infected individual: less than 20%), while for shorter-duration travel (e.g. 1-2 days), a 5-day	Measure versus no measure	Positive (▲)

(Continued)

	quarantine will prevent 50-60% of further transmission (probability of releasing an infected individual: 40-50%). A 3-day quarantine after longer trips will prevent 70-75% of further transmission (probability of releasing an infected individual: 25-30%), while for shorter trips only 15-25% (probability of releasing in infected individual: 75-85%). For shorter quarantine periods to accomplish a similar impact, greater compliance is necessary.		
	Clifford 2020b: Compared with symptom/exposure-based entry screening alone, requiring a mandatory quarantine in the UK for travellers from the EU would reduce the risk of releasing infected travellers into the community, with the magnitude of the effect dependent on the length of quarantine (3-day quarantine: RR 0.59, 95% CI 0.28 to 0.85); 14-day quarantine: RR 0.00, 95% CI 0.00 to 0.01)). Requiring a pre-flight test only would slightly improved the effectiveness of the quarantine of travellers.	Measure versus alternative measure	Positive (▲)
	Steyn 2020: Assuming a moderate risk of transmission within quarantine settings in New Zealand, the probability of releasing an infected individual into the community would be 4% for a 14-day quarantine. This is lower than the probability of 50% of releasing an infected individual for exit and entry testing of travellers only. Assuming no risk of transmission within quarantine facilities, a 14-day quarantine would yield a probability of 0% of releasing an infected individual into the community compared with 48% for exit and entry testing only. Assuming a high risk of transmission in quarantine facilities, the probability of releasing an infected individual into the community would be 28% compared with 53% for entry and exit testing only.	Measure versus alternative measure	Positive (▲)

Appendix 14. Quarantine and screening at borders: study-by-study overview of the evidence contributing to each outcome (modelling studies)

Outcome	Number of studies	Overview of effect by study	Comparison used in each study	Effect direction per study (positive ▲; negative ▼; no change/mixed effects/conflicting findings ◀▶)
Outcome category: cases avoided due to the measure				
No contributing studies				
Outcome category: shift in epidemic development				

(Continued)

Time to outbreak	1 modelling study	Wilson 2020: This study assessed the effect of combinations of travel-related control measures on the time to outbreak in a hypothetical disease-free area (modelled on New Zealand). Compared with exit screening alone (2.2. years to outbreak (95% CI 0.06 to 8.11), various co-interventions, such as entry screening, quarantine, PCR testing of incoming travellers, in-flight wearing of masks, contact tracing and self-reporting of symptoms would lead to delays of the outbreak. For example, assuming one flight per day (7.1% of normal travel volume) and 50% sensitivity, exit screening, in-flight wearing of masks and entry screening would lead to 3.5 years to outbreak (95% CI 0.09 to 12.9); exit screening and in-flight wearing of masks would lead to 3.3 years to outbreak (95% CI 0.08 to 12.1), exit screening, in-flight wearing of masks, entry screening and a 7-day quarantine 5.8 years (95% CI 0.15 to 21.5), and exit screening, in-flight wearing of masks, and PCR testing on day 1 4.4 years (95% CI 0.11 to 16.1). The most effective combination of measures was found to be exit screening, in-flight wearing of masks, entry screening and a 14-day quarantine of all arriving travellers, yielding 34.1 years to outbreak (95% CI 0.86 to 126).	Combined measures versus single measure	Positive (▲)

Outcome category: cases detected due to measure

Days at risk of transmission	2 modelling studies	Clifford 2020b: Compared with a single PCR test upon arrival in the UK from EU countries (2.0 days at risk (95% CI 0 to 10.8)), requiring all travellers to quarantine before being tested would lead to further reductions in the days that the travellers remain at risk of transmitting the infection into the community. The effect was shown to increase with longer periods of quarantine before testing (3-day quarantine: 0.4 days at risk (95% CI 0 to 10.2); 14-day quarantine: 0 days at risk (95% CI 0 to 0)). Requiring two PCR tests during the quarantine period (2, 4, or 6 days after the first test) was slightly better, yet largely comparable to a single PCR test and quarantine (3-day quarantine: 0 days at risk (95% CI 0 to 8.9); 9-day quarantine: 0 days (95% CI 0 to 3.1)).	Combined measures versus single measure	Positive (▲)
		Russell WA 2020: Compared with a quarantine of all incoming travellers at a hypothetical point of entry (2-day quarantine: 1.8 days at risk (95% CI 1.6 to 2.2); 14-day quarantine: 0.53 days at risk (95% CI 0.46 to 0.60)), additional testing upon arrival and 24 hours before the end of quarantine would lead to reductions in days at risk (2-day quarantine: 1.5 days at risk (95% CI 1.3 to 1.8); 14-day quarantine: 0.52 days at risk (95% CI 0.46 to 0.59).	Combined measures versus single measure	Positive (▲)
Proportion of cases detected	2 modelling studies	Bays 2020: Compared with entry screening alone, which would detect 0.8% of cases, entry screening followed by an isolation period and a test at the end of the isolation period would increase the proportion of cases detected in a hypothetical popula-	Combined measures versus single measure	Positive (▲)

(Continued)

		tion: a shorter isolation period of 3 days would detect between 41 and 62% of cases (depending on travel/flight duration), while a longer isolation period of 14 days would detect almost all cases (99% in all scenarios of travel/flight duration).		
		Taylor 2020: Compared with 7-day self-isolation only, which would detect 51% (95% CI 47 to 56) of cases for a 7-day isolation period, requiring self-isolation and subsequent testing all travellers arriving in the UK would increase the proportion of cases detected. The magnitude of effect would increase with the number of days in self-isolation: testing 4 days and 7 days after arrival would allow detection of 64.3% and 74.3% of infected individuals, respectively. An additional test upon arrival would improve the proportion of cases detected only very slightly.	Combined measures versus single measure	Positive (▲)
Probability of releasing an infected individual into the community	3 modelling studies	Ashcroft 2020: This study assesses the effect of combining quarantine with testing on different days at a hypothetical point of entry on further transmission. Assuming travel duration of 7 days and a test result delay of 2 days, testing on arrival (i.e. release on day 2) would prevent 54% of further transmission (probability of releasing an infected individual: 46%). Testing on day 3 (i.e. release on day 5) and testing on day 5 (i.e. release on day 7) would prevent approximately 90% and 99% of further transmission, respectively (probability of releasing an infected case: 10% and 1%, respectively). The proportion of further transmission prevented by quarantine alone is lower: a quarantine period of 2 days, 5 days and 7 days would prevent 40%, 82% and 95% of further transmission, respectively (probability of releasing an infected case: 60%, 18% and 5%).	Combined measures versus single measure	Positive (▲)
		Clifford 2020b: Compared to PCR testing upon arrival in the UK from EU countries (RR 0.55, 95% CI 0.28 to 0.83), with symptom screening upon arrival as reference), requiring all travellers to quarantine before being tested would reduce the risk of releasing an infected individual into the community. The effect would depend on the length of the quarantine period: risk ratio for a 3-day quarantine: 0.22 (95% CI 0.02 to 0.48); risk ratio for a 14-day quarantine: 0.01 (95% CI 0.00 to 0.03). Requiring two PCR tests during the quarantine period and before release was slightly better, yet largely comparable to a single PCR test and quarantine (risk ratio for a 3-day quarantine: 0.17 (95% CI 0.00 to 0.39); risk ratio for a 9-day quarantine: 0.01 (95% CI 0.00 to 0.11)).	Combined measures versus single measure	Positive (▲)
		Steyn 2020: For a 14-day quarantine with two tests on days 3 and 12, and a moderate risk of transmission within the quarantine facility, the probability of releasing an infected individual into the community was 2%; this is lower than both 14-day quarantine only (4%) and departure and arrival testing	Combined measures versus single measure	Mixed (◀▶)

(Continued)

only (50%). The probability of releasing an infected individual would decrease further with no risk of transmission in quarantine (0%), while it would increase with a high risk of transmission (7%).

WHAT'S NEW

Date	Event	Description
26 February 2021	New citation required but conclusions have not changed	Our conclusions remain unchanged.
26 February 2021	New search has been performed	We updated the searches. In this update we removed all studies on SARS and MERS and only included COVID-19 studies.

HISTORY

Review first published: Issue 9, 2020

Date	Event	Description
17 September 2020	Amended	Minor typographical error corrected

CONTRIBUTIONS OF AUTHORS

JB, AM and ER developed the protocol, and JB and AM co-ordinated the overall review process. JS had substantial intellectual contribution in the conduct of this review, including data extraction, risk of bias assessment and evidence synthesis. JB, AM, and EAR put together the report. All other review authors are listed in alphabetical order, with everyone making substantial intellectual contribution.

Protocol development: JB, AM and ER

Citation searches: AM, SV

Title and abstract screening: RB, MC, KG, CK, SK, JS, KS, BV, SV, KW

Full-text screening: JB, RB, MC, KG, CK, SK, AM, JS, KS, BV, SV, KW

Data extraction: JB, KMFE-F, CK, LMP, AM, SN, PvP, JMS, JS, BV, SV, KW

Risk of bias/quality assessment: JB, KMFE-F, SH, ML, TL, JMS

GRADE assessment: JB, AM, ER

Evidence synthesis: JB, AM, ER, JMS

Manuscript preparation: JB, RB, MC, KMFE-F, KG, SH, OH, CK, ML, TL, AM, LMP, PvP, ER, JMS, JS, KS, BV, SV, KW.

DECLARATIONS OF INTEREST

Jacob Burns: the Chair for Public Health and Health Services Research at the Institute for Medical Information Processing, Biometry and Epidemiology received funding from the WHO for the conduct of this review. The WHO provided input which informed the review protocol and scope. The WHO did not exert any influence on how findings were interpreted. The role of the WHO is clearly specified in the review itself. The Chair for Public Health and Health Services Research at the Institute for Medical Information Processing, Biometry

Susan Norris: the Chair for Public Health and Health Services Research at the Institute for Medical Information Processing, Biometry and Epidemiology received funding from the WHO for the conduct of this review. The WHO provided input which informed the review protocol and scope. The WHO did not exert any influence on how findings were interpreted. The role of the WHO is clearly specified in the review itself. The Chair for Public Health and Health Services Research at the Institute for Medical Information Processing, Biometry and Epidemiology received funding from the German Federal Ministry of Education and Research (BMBF) as part of the COVID-19 evidence ecosystem (CEOsys) project. No other conflicts of interest are known.

Lisa M Pfadenhauer: the Chair for Public Health and Health Services Research at the Institute for Medical Information Processing, Biometry and Epidemiology received funding from the WHO for the conduct of this review. The WHO provided input which informed the review protocol and scope. The WHO did not exert any influence on how findings were interpreted. The role of the WHO is clearly specified in the review itself. The Chair for Public Health and Health Services Research at the Institute for Medical Information Processing, Biometry and Epidemiology received funding from the German Federal Ministry of Education and Research (BMBF) as part of the COVID-19 evidence ecosystem (CEOsys) project. No other conflicts of interest are known.

Peter von Philipsborn: the Chair for Public Health and Health Services Research at the Institute for Medical Information Processing, Biometry and Epidemiology received funding from the WHO for the conduct of this review. The WHO provided input which informed the review protocol and scope. The WHO did not exert any influence on how findings were interpreted. The role of the WHO is clearly specified in the review itself. The Chair for Public Health and Health Services Research at the Institute for Medical Information Processing, Biometry and Epidemiology received funding from the German Federal Ministry of Education and Research (BMBF) as part of the COVID-19 evidence ecosystem (CEOsys) project. No other conflicts of interest are known.

Kerstin Sell: the Chair for Public Health and Health Services Research at the Institute for Medical Information Processing, Biometry and Epidemiology received funding from the WHO for the conduct of this review. The WHO provided input which informed the review protocol and scope. The WHO did not exert any influence on how findings were interpreted. The role of the WHO is clearly specified in the review itself. The Chair for Public Health and Health Services Research at the Institute for Medical Information Processing, Biometry and Epidemiology received funding from the German Federal Ministry of Education and Research (BMBF) as part of the COVID-19 evidence ecosystem (CEOsys) project. No other conflicts of interest are known.

Julia Stadelmaier: the Institute for Evidence in Medicine at the University of Freiburg received funding from the German Federal Ministry of Education and Research (BMBF) as part of the COVID-19 evidence ecosystem (CEOsys) project. No other conflicts of interest are known.

Ben Verboom: the Chair for Public Health and Health Services Research at the Institute for Medical Information Processing, Biometry and Epidemiology received funding from the WHO for the conduct of this review. The WHO provided input which informed the review protocol and scope. The WHO did not exert any influence on how findings were interpreted. The role of the WHO is clearly specified in the review itself. The Chair for Public Health and Health Services Research at the Institute for Medical Information Processing, Biometry and Epidemiology received funding from the German Federal Ministry of Education and Research (BMBF) as part of the COVID-19 evidence ecosystem (CEOsys) project. No other conflicts of interest are known.

Stephan Voss: the Chair for Public Health and Health Services Research at the Institute for Medical Information Processing, Biometry and Epidemiology received funding from the WHO for the conduct of this review. The WHO provided input which informed the review protocol and scope. The WHO did not exert any influence on how findings were interpreted. The role of the WHO is clearly specified in the review itself. The Chair for Public Health and Health Services Research at the Institute for Medical Information Processing, Biometry and Epidemiology received funding from the German Federal Ministry of Education and Research (BMBF) as part of the COVID-19 evidence ecosystem (CEOsys) project. No other conflicts of interest are known.

Katherina Wabnitz: the Chair for Public Health and Health Services Research at the Institute for Medical Information Processing, Biometry and Epidemiology received funding from the WHO for the conduct of this review. The WHO provided input which informed the review protocol and scope. The WHO did not exert any influence on how findings were interpreted. The role of the WHO is clearly specified in the review itself. The Chair for Public Health and Health Services Research at the Institute for Medical Information Processing, Biometry and Epidemiology received funding from the German Federal Ministry of Education and Research (BMBF) as part of the COVID-19 evidence ecosystem (CEOsys) project. No other conflicts of interest are known.

Eva Rehfuess: the Chair for Public Health and Health Services Research at the Institute for Medical Information Processing, Biometry and Epidemiology received funding from the WHO for the conduct of this review. The WHO provided input which informed the review protocol and scope. The WHO did not exert any influence on how findings were interpreted. The role of the WHO is clearly specified in the review itself. The Chair for Public Health and Health Services Research at the Institute for Medical Information Processing, Biometry and Epidemiology received funding from the German Federal Ministry of Education and Research (BMBF) as part of the COVID-19 evidence ecosystem (CEOsys) project. No other conflicts of interest are known.

SOURCES OF SUPPORT

Internal sources

- No sources of support supplied

External sources

- World Health Organization (WHO), Other

  Some of this article is based on a study commissioned and paid for by the WHO, who provided input which informed the review protocol and scope.
- German Ministry of Education and Research, Germany

  This review was undertaken in the context of COVID-19 Evidence Ecosystem project, funded by the German Ministry of Education and Research.

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

There are some differences between the protocol in Appendix 1 and some sections of the review.

Title: based on peer-review feedback, we have revised the title for this review to add the word "international" to explicate the focus of the review. While this is in line with the scope of our protocol and review, we had not specified this in the title of the protocol.

Criteria for considering studies for this review: for this review, we added one post-hoc exclusion criterion which we had not specified in the review protocol. This relates to studies of interventions assessing international travel but not concerned with cross-border impacts, i.e. interventions to contain transmission within closed populations and only assessing their effect on those closed populations (e.g. on cruise ships or within detention centres). Furthermore, for this review update, we changed the eligibility criteria to only include studies focusing on SARS-CoV-2/COVID-19 and exclude those focusing on SARS-CoV-1/SARS and MERS-CoV/MERS. The latter, however, was considered in the first version of the review.

Secondary outcomes: for this review, we have added an additional secondary outcome, namely, user acceptability, based on exchanges with the WHO. This was not specified in the review protocol.

Data extraction and management: in the review protocol we specified that we would search for data from external sources, such as daily COVID-19 situation reports published by the WHO, to enhance our understanding of the specific features of travel-related control measures and the stage of pandemic at the time of implementing the measure. However, because of the lack of consistency of the information in these sources (e.g. discrepancies in how the WHO reports describe the stage of pandemic in earlier months of 2020), we did not search these sources in the review.

Assessment of risk of bias in included studies: in the review protocol we specified using multiple 'Risk of bias' tools to assess the broad range of study types in the review, namely RoB 2 tool for experimental studies, Risk of Bias in Non-randomised Studies – of Interventions (ROBINS-I) for quasi-experimental and observational studies and a bespoke tool we developed for assessing the quality of modelling studies. However, given that we identified additional sets of papers evaluating entry and/or exit screening that were more related to diagnostic studies than intervention evaluation, we decided post-protocol to apply an additional 'Risk of bias' tool, namely the Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) to assess risk of bias of these studies. Furthermore, based on the rounds of feedback from the modelling experts in our team, we slightly adapted the criteria in the bespoke tool we developed for assessing the quality of modelling studies. These included criteria related to input data, internal and external validity. However, this did not change the overall structure of the tool.

Data synthesis: in our review we identified several studies which met the predefined inclusion criteria, which however did not provide directly policy-relevant evidence. These included (i) observational studies evaluating entry and/or exit screening measures reporting only limited data regarding the effectiveness of the measure, (ii) observational ecological studies examining the aggregated impact of various travel-related control measures across countries, and (iii) modelling studies using overly simplistic or theoretical assumptions and presenting abstract findings. We classified these studies as 'supporting studies' and summarised their results descriptively in an appendix. This process, however, was not prespecified in our review protocol.

Subgroup analyses and investigation of heterogeneity: we specified in our protocol that we would investigate the influence of several sources of heterogeneity on the impact of interventions, including the stage of the pandemic during which the intervention was implemented, whether the country is an island state or not, assumed infectious disease parameters, and whether a single or a package of control measures were implemented. While we reflect on the potential impact of some of these sources in our review, we did not explore heterogeneity as planned, because of lack of data on these sources.

Sensitivity analyses: in our review protocol we specified that we would conduct a sensitivity analysis to assess the impact of studies at high risk of bias on the review findings. However, we did not conduct sensitivity analysis in the review given the nature of our narrative synthesis and as most of the evidence base was comprised of modelling studies with major quality concerns.

INDEX TERMS

Medical Subject Headings (MeSH)

Bias; Communicable Diseases, Imported [epidemiology] [prevention & control]; COVID-19 [epidemiology] [*prevention & control]; Internationality; Models, Theoretical; Observational Studies as Topic; Pandemics [*prevention & control]; Quarantine; *SARS-CoV-2; *Travel-Related Illness

MeSH check words

Humans