Author for correspondence:
Wan Yang
e-mail: wy2202@cumc.columbia.edu

Reconciling the efficacy and effectiveness of masking on epidemic outcomes

Wan Yang[1,4] and Jeffrey Shaman[2,3]

[1]Department of Epidemiology, Mailman School of Public Health, [2]Department of Environmental Health Sciences, Mailman School of Public Health, [3]Columbia Climate School, Columbia University, New York, NY, USA, and [4]Herbert Irving Comprehensive Cancer Center, Columbia University Medical Center, New York, NY, USA

WY, 0000-0002-7555-9728

During the COVID-19 pandemic, mask wearing in public settings has been a key control measure. However, the reported effectiveness of masking has been much lower than laboratory measures of efficacy, leading to doubts on the utility of masking. Here, we develop an agent-based model that comprehensively accounts for individual masking behaviours and infectious disease dynamics, and test the impact of masking on epidemic outcomes. Using realistic inputs of mask efficacy and contact data at the individual level, the model reproduces the lower effectiveness as reported in randomized controlled trials. Model results demonstrate that transmission within households, where masks are rarely used, can substantially lower effectiveness, and reveal the interaction of nonlinear epidemic dynamics, control measures and potential measurement biases. Overall, model results show that, at the individual level, consistent masking can reduce the risk of first infection and, over time, reduce the frequency of repeated infection. At the population level, masking can provide direct protection to mask wearers, as well as indirect protection to non-wearers, collectively reducing epidemic intensity. These findings suggest it is prudent for individuals to use masks during an epidemic, and for policymakers to recognize the less-than-ideal effectiveness of masking when devising public health interventions.

1. Introduction

During the coronavirus disease 2019 (COVID-19) pandemic, universal mask wearing in public settings has been implemented in many places throughout the world as a means of mitigating the pandemic. A wide variety of masks, ranging from cotton masks to N95 respirators, have been used by the general population throughout the pandemic. Intuitively, masks create a physical barrier between the wearer's airway and the environment, reducing the chance of both disseminating and acquiring infection. In addition, a number of laboratory experiments, often based on manikin simulators, have demonstrated high efficacies of different masks including cotton masks; measured reductions of inhaled particles in laboratory settings ranged from approximately 50% to greater than 90% [1–3]. However, the reported effectiveness of masking, demonstrated from randomized controlled trials (RCTs), has been much lower than the laboratory experiments, ranging from 10 to 25% [4,5] to inconclusive in a 2023 Cochrane review [6]. Further, many places have endured multiple COVID-19 pandemic waves since March 2020, despite mask mandates implemented in public settings. The large discrepancy between the *efficacy* (measured under ideal conditions of mask quality and correct use, e.g. in laboratories measured by the filtration rate) and *effectiveness* (measured under real-world, often imperfect conditions, e.g. in RCTs) of masking has prompted debates on the need for mask use in the general population [7,8].

To explain this large discrepancy, studies have cited a number of factors. These mainly include compliance (e.g. the percentage of people actually

wearing masks could be lower than the expected 100% among the test arm of an RCT), consistency (e.g. participants may not wear masks all the time, including the moment of exposure), and timing of implementation (e.g. there may be little circulation/presence of pathogen during an RCT to support demonstration of an effect) [8–10]. However, it is difficult to test these factors in real-world settings and, as such, opinions and practices are divided—some dismiss the low effectiveness estimates from RCTs and continue to mask; others dismiss the high laboratory efficacies and refuse to mask. In addition, it remains unclear when and to what extent masking can be effective at the population level.

Here, to inform more effective individual preventive care and public health measures, we develop an agent-based model that comprehensively accounts for individual masking behaviours and efficacies as well as the nonlinear transmission process, and test the combined impact of these processes at the population level. Using realistic inputs of mask efficacy (i.e. 50–80%, mean = 66%; electronic supplementary material, table S1) and contact data at the individual level, the model is able to reproduce the lower effectiveness (e.g. less than 25%) at the population level, as observed in real-world settings. Model results highlight the main factors contributing to the lower effectiveness, illustrate the complexity of measuring the underlying impact of masking due to nonlinear epidemic dynamics, and provide a more complete picture of the protective impact of masking (e.g. reducing repeated infections).

2. Results

2.1. Overview of the agent-based model and analyses

The model used here has several key features. First, it explicitly simulates the transmission process to capture the contagious nature of infectious diseases. That is, individual risk of infection is determined by behaviour (e.g. contact rate and mask preference) as well as the people with whom an individual interacts (e.g. the infection status of contacts). This structure depicts the connectivity and spread of infection across a population, which is probably a more realistic representation of the real world (versus e.g. regression models). Second, the model separates household and non-household contact and transmission. This formulation not only allows modelling of the relative importance of household and non-household transmission based on contact data, but also allows modelling of specific masking behaviour in each setting (e.g. masks are often not used within a household). Third, the model records the history of each infection and vaccination to account for immune protection and waning for each individual and, through the aforementioned inter-person interactions, the combined outcomes at the population level. Fourth, the model is age-specific, which allows age-specific contact rates (e.g. higher among school-age children), efficacy of masking (e.g. higher for older adults) and disease severity (electronic supplementary material, table S1). This differentiation allows more realistic representation of population heterogeneity. Fifth, for a given mask type with known filtration efficiency, wearers may have a higher chance of infection when encountering a pathogen with a lower median infectious dose (ID_{50}) than that with a higher ID_{50}, i.e. akin to having a lower mask efficacy. While we do not explicitly model pathogen-specific characteristics, the wide range of masking efficacies explored in our model simulations (electronic supplementary material, table S1) allows examination of infection outcomes (i.e. mask effectiveness) when mask efficacy is varied due to e.g. different IDs. Lastly, the model is run continuously over multiple years, where an individual is subject to varying background (environmental) infection risks as the population-level infection prevalence varies over time. As such, it represents mask effectiveness and the cumulative impact for different background (environmental) infection risks. Taken together, unlike previous approaches, we more comprehensively capture individual-level effects, nonlinear epidemic dynamics, and overall impact at the population level.

Using the model, we simulated a COVID-19-like pandemic and subsequent epidemic waves in a population of 100 000 people, from 1 March 2020 (roughly the pandemic onset) to 31 December 2024 (roughly a 5-year period). The basic reproduction number (R_0, i.e. the average number of secondary infections in a naive population) was set to 2.6—similar to estimates for the ancestral SARS-CoV-2 [11–14] that initiated the COVID-19 pandemic—in our main analysis and counterfactual analysis using contact data from the POLYMOD study [15] and assuming an infection probability of 50% (10%) for physical (nonphysical) contact. To simulate pandemic dynamics, we used mobility data to adjust contact rates and vaccination data to account for vaccine protection, for all simulations. To test the impact of masking, we varied the masking rate (i.e. the fraction of population using masks, 10–100%) and the timing of masking (always wear masks outside the home versus when prevalence is above 0.01% to 2%) and compared simulated epidemic outcomes with a baseline setting without masking.

We compute the masking effectiveness at the population level as the reduction in a given health outcome (e.g. the total number of infections in a population) under a given masking setting compared with the baseline setting without masking (see Methods). In doing so, the population-specific characteristics are largely cancelled out, such that the effectiveness estimates would be more generalizable to different populations. This measure is also akin to those reported in cluster RCTs, which typically compare the overall attack rates in communities assigned to wear masks (where there may be different levels of adherence, similar to the simulations here under different masking settings) with communities not assigned to wear masks (similar to our baseline setting without masking). This measure is also consistent with non-cluster RCTs reporting reductions in the attack rate of a test arm with masking compared with a control arm without masking. In addition, as noted in the Introduction, reported masking effectiveness measures from various RCTs fall within similar ranges [4–6]. As such, here we compare our model results with reported estimates in RCTs overall (i.e. without separating individual types of RCTs).

2.2. The model recapitulates masking effectiveness reported in randomized controlled trials

Figure 1 shows example simulated daily infection rates during the approximately 5-year study period. Under the most vigilant scenario in which people choosing to wear masks (i.e. wearers) always do so outside the home, increasing the masking rate greatly reduces epidemic size; this reduction is consistent across all years (figure 1a). However, even with a near 100% masking rate (i.e. all
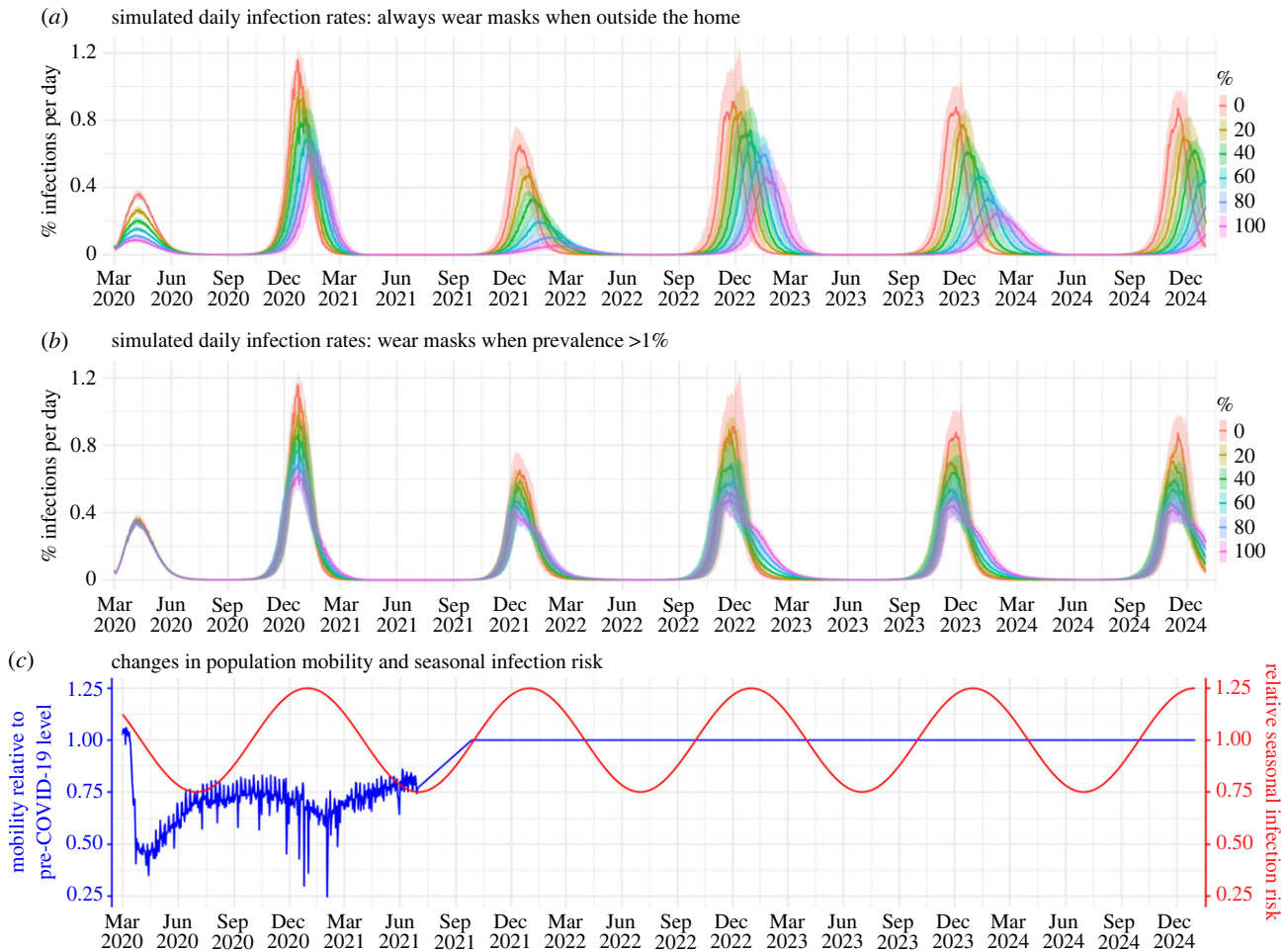
Figure 1. Example simulated epidemic infection rates. For these simulations, the model assumes an R_0 of 2.6 and no mask wearing at home. Coloured lines (median) and shaded areas (interquartile ranges) show daily infection rates for simulated populations with different masking rates (see legend) and timing of masking: (*a*) mask wearers do so outside the home at all times or (*b*) only when infection prevalence is greater than 1%. For context, (*c*) shows changes in daily population mobility (representing reductions in contact rate due to social distancing rules) and seasonal infection risk (representing effects due to environmental conditions), used in the simulations. Mobility data for New York City during March 2020–June 2021 (shown here) were used directly as model inputs. From July 2021 onward, we assumed there was no social distancing and increased the mobility linearly to the pre-COVID level over a three-month period and then remained at that level (i.e. set to 1). Trends in seasonal infection risk were represented using a sinusoidal function (see details in Methods).

except for those aged less than 1 year use masks), epidemic waves, albeit of much smaller size, persist. In addition, the effectiveness decreases substantially, when masking is implemented only when population prevalence of infection is high (e.g. greater than 1% in figure 1*b*). It is further evident from figure 2*a*, across a wide range of scenarios and time frames, that there are many instances that estimated effectiveness (measured as the reduction in total number of infections compared with the baseline without masking) is less than 25%—the higher bound of estimates reported in RCTs—including instances with high masking rates (e.g. greater than 80%) if masking is enacted late (see cells within the black polygons in figure 2*a*).

2.3. Factors contributing to lower masking effectiveness in the real world

To understand the main factors behind low masking effectiveness, we hypothesize that transmission may persist in household settings where contact tends to be greater, yet masking is rare. As shown in figure 3*a* (contact rates formulated using the POLY-MOD survey data [15]), for most individuals, household contact accounts for the majority of total contact, and are hence a likely principal source of infection. Thus, to test the impact of household transmission, we simulated a counterfactual setting in which people also wear masks at home (figure 3*b*).

Three main findings emerge from the counterfactual analysis. First, compared with when masking is solely applied outside the home (main setting, figure 2*a*), the effectiveness would be much greater should masking be also universally applied at home (counterfactual, figure 3*b*). For instance, assuming a 50% masking rate and masking at all times, over the 5-year study period, the counterfactual effectiveness is 62%, more than twice of the estimate in the main setting (25%; figure 2*a*, last panel, the cell in the first column and middle row). Second, it is evident from figure 3*b* that the effectiveness is higher than the expected direct effect should masking only affect the wearers. For instance, with a 50% masking rate and masking at all times, the expected effectiveness is 33% (i.e. 50% × 66% mean efficacy), much lower than the 62% estimate. This greater effectiveness highlights the indirect protection of masking to the entire population including non-wearers, an effect that is hidden in real-world settings. Third, there is often a delay in awareness of disease prevalence and, in turn, a delay in masking. Such delays further facilitate introduction of infection into households and onward transmission (both within and outside households), leading to decreases in the overall
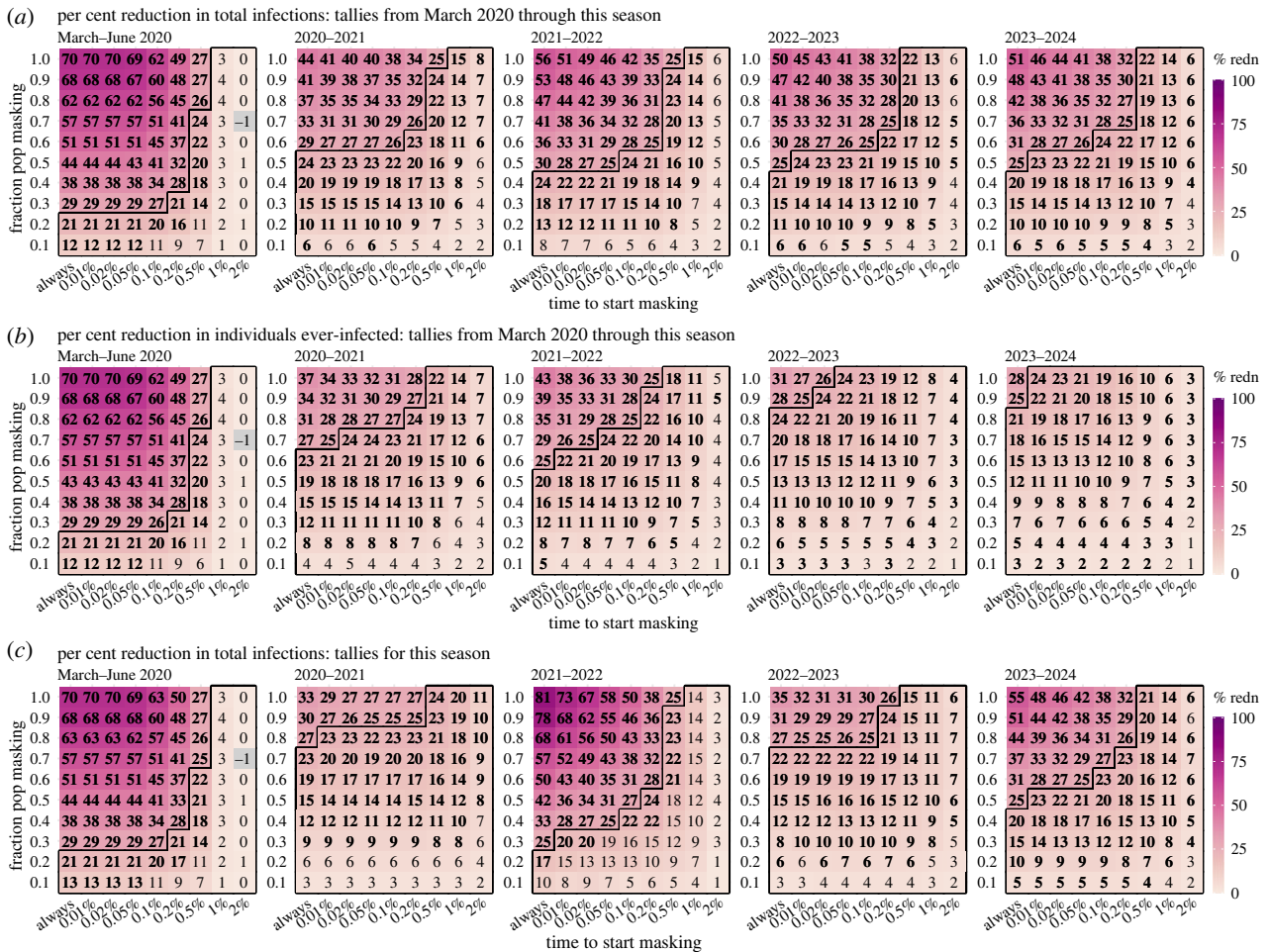
Figure 2. Estimated effectiveness of masking in reducing infection rates at the population level (main analysis assuming an R_0 of 2.6 and no masking at home). Heatmaps show (a) per cent reduction of cumulative total infections from March 2020 through different seasons (see panel subtitles, same below), (b) per cent reduction of the cumulative fraction of population ever-infected, i.e. regardless of number of (re)infections, and (c) per cent reduction of total infections tallied for each season. Here, a season runs 1 year from 1 July to 30 June of the following year. Numbers in the cells show median effectiveness estimates (computed per equation (4.2)) for the corresponding masking rate (y-axis) and timing of masking (x-axis, percentages show population infection prevalence when masking is enacted). Bold fonts indicate the 2.5th % of the estimate is above 0 (i.e. significant at $a = 0.05$ level); note the 95% intervals here are probably narrower than measures in RCTs as the model assumes all infections are identified for the entire population. Negative values (i.e. increases; see grey cells) are possible due to model stochasticity (N.B., none are significant at $a = 0.05$ level). The black lines delineate cells with a median effectiveness estimate less than 25%.

effectiveness over time. For instance, assuming a 50% masking rate and masking at all times, the estimated effectiveness in the main setting decreased from 44% during March–June 2020 to 25% at the end of the 5-year study period (figure 2a, first versus last panel), whereas the counterfactual estimates remained above 60% over the entire 5-year study period (figure 3b, first versus last panel).

The model results also illustrate how system nonlinear dynamics can affect masking effectiveness. For example, during the first wave (March–June 2020), estimated effectiveness surpasses the expected, under most combinations of masking rate and timing (figure 2a, the number in the cell > masking rate × 66% mean efficacy). This higher-than-expected effectiveness comes from the transmission reduction due to masking, which, in combination with the reduction in contact rate and less conducive environmental conditions in late spring/summer (figure 1c), reduces the effective reproduction number R_e to less than 1, preventing a large pandemic wave (note: R_e measures the average number of secondary infections and unity is the epidemic threshold). Conversely, when R_e is much higher than 1, the effectiveness tends to be lower than expected due to 'leakage' via household transmission, as noted above. For instance, a large wave occurred during the 2020–2021 season (see figure 1a, second wave) under all masking rates, probably due to high population susceptibility and relaxed implementation of other non-pharmaceutical interventions (NPIs; e.g. social distancing; figure 1c, blue line) coinciding with more conducive transmission conditions during wintertime (figure 1c, red line). As a result, the reduction in epidemic magnitude was much smaller than during the first wave. This is also evident in the cumulative effectiveness over time (figure 2a; compare the higher effectiveness estimates for March–June 2020 in the first panel with the lower effectiveness estimates for subsequent seasons in remaining panels).

2.4. Measurement issues can further affect reported masking effectiveness

In addition to the factors noted above, the method and timing of measurement could both affect the reported effectiveness of masking. For the former, reported effectiveness would be lower than actual, if repeated infections are not counted. For instance, the number of people ever infected will inevitably increase over time. Studies that only measure the ever-infected while omitting
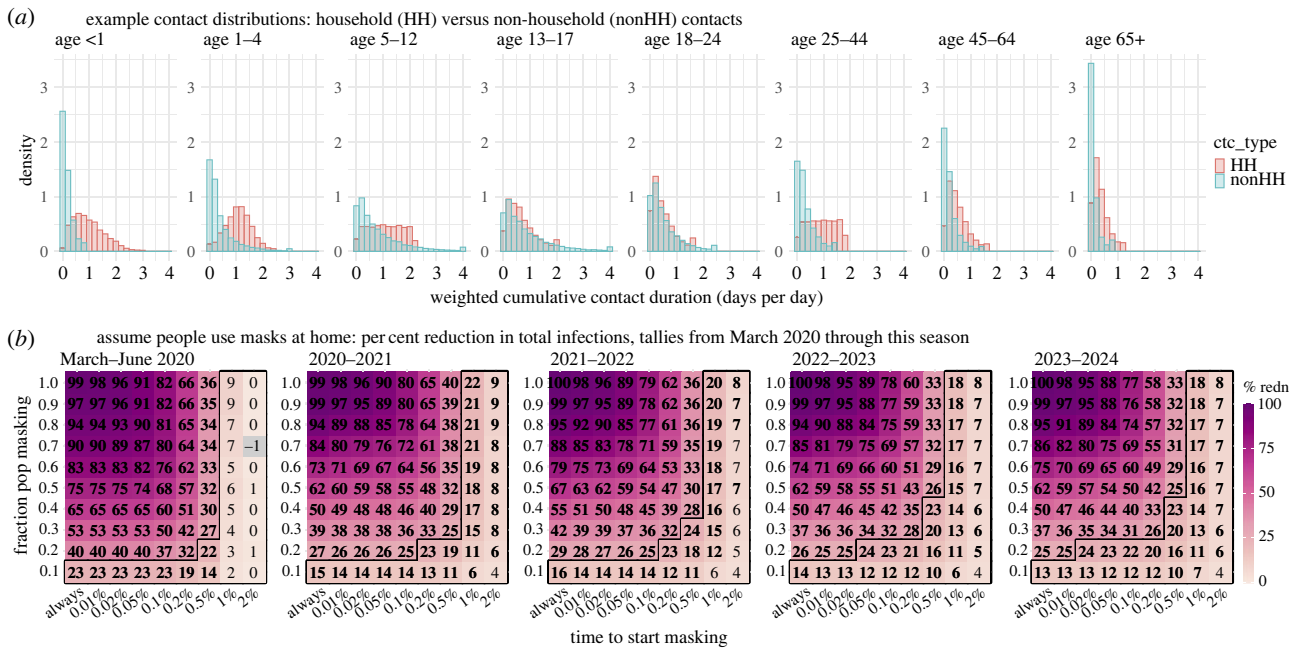
(a) example contact distributions: household (HH) versus non-household (nonHH) contacts

age <1 · age 1–4 · age 5–12 · age 13–17 · age 18–24 · age 25–44 · age 45–64 · age 65+

y-axis: density · *x-axis:* weighted cumulative contact duration (days per day)

ctc_type: HH, nonHH

(b) assume people use masks at home: per cent reduction in total infections, tallies from March 2020 through this season

y-axis: fraction pop masking · *x-axis:* time to start masking · legend: % redn (100, 75, 50, 25, 0)

March–June 2020

fraction pop masking	always	0.01%	0.02%	0.05%	0.1%	0.2%	0.5%	1%	2%
1.0	99	98	96	91	82	66	36	9	0
0.9	97	97	96	91	82	66	35	9	0
0.8	94	94	93	90	81	65	34	7	0
0.7	90	90	89	87	80	64	34	7	−1
0.6	83	83	83	82	76	62	33	5	0
0.5	75	75	75	74	68	57	32	6	1
0.4	65	65	65	65	60	51	30	5	0
0.3	53	53	53	53	50	42	27	4	0
0.2	40	40	40	40	37	32	22	3	1
0.1	23	23	23	23	23	19	14	2	0

2020–2021

fraction pop masking	always	0.01%	0.02%	0.05%	0.1%	0.2%	0.5%	1%	2%
1.0	99	98	96	90	80	65	40	22	9
0.9	99	97	95	89	80	65	39	21	9
0.8	94	89	88	85	78	64	38	21	9
0.7	84	80	79	76	72	61	38	21	8
0.6	73	71	69	67	64	56	35	19	8
0.5	62	60	59	58	55	48	32	18	8
0.4	50	49	48	48	46	40	29	17	8
0.3	39	38	38	38	36	33	25	15	8
0.2	27	26	26	26	25	23	19	11	6
0.1	15	14	14	14	14	13	11	6	4

2021–2022

fraction pop masking	always	0.01%	0.02%	0.05%	0.1%	0.2%	0.5%	1%	2%
1.0	100	98	96	89	79	62	36	20	8
0.9	99	97	95	89	78	62	36	20	7
0.8	95	92	90	85	77	61	36	19	7
0.7	88	85	83	78	71	59	35	19	7
0.6	79	75	73	69	64	53	33	18	7
0.5	67	63	62	59	54	47	30	17	7
0.4	55	51	50	48	45	39	28	16	6
0.3	42	39	39	37	36	32	24	15	6
0.2	29	28	27	26	25	23	18	12	5
0.1	16	14	14	14	14	12	11	6	4

2022–2023

fraction pop masking	always	0.01%	0.02%	0.05%	0.1%	0.2%	0.5%	1%	2%
1.0	100	98	95	89	78	60	33	18	8
0.9	99	97	95	88	77	59	33	18	7
0.8	94	90	88	84	75	59	33	17	7
0.7	85	81	79	75	69	57	32	17	7
0.6	74	71	69	66	60	51	29	16	7
0.5	62	59	57	54	50	42	25	16	7
0.4	50	47	46	45	42	35	23	14	7
0.3	37	36	35	34	32	28	20	13	6
0.2	26	25	25	24	23	21	16	11	5
0.1	14	13	13	12	12	12	10	6	4

2023–2024

fraction pop masking	always	0.01%	0.02%	0.05%	0.1%	0.2%	0.5%	1%	2%
1.0	100	98	95	88	77	58	33	18	8
0.9	99	97	95	88	76	58	32	18	7
0.8	95	91	89	84	74	57	32	17	7
0.7	86	82	80	75	69	55	31	17	7
0.6	75	70	69	65	60	49	29	16	7
0.5	62	59	57	54	50	42	25	16	7
0.4	50	47	46	44	40	33	23	14	7
0.3	37	36	35	34	31	26	20	13	6
0.2	25	25	24	23	22	20	16	11	6
0.1	13	13	13	12	12	12	10	7	4

Figure 3. The impact of household transmission. For these simulations, the model assumes an R_0 of 2.6 (i.e. the same as in the main analysis). (a) Example distributions of contact duration, formulated based on the POLYMOD survey (see details in Methods). (b) Counterfactual estimates of masking effectiveness in reducing total number of infections, assuming people universally wear masks at home. Numbers in the cells show the median effectiveness estimates (computed per equation (4.2)) for the corresponding masking rate (y-axis) and timing of masking (x-axis, percentages show population infection prevalence when masking is enacted). Bold fonts indicate the 2.5th% of the estimate is above 0 (i.e. significant at $a = 0.05$ level); note the 95% intervals here are probably narrower than measures in RCTs as the model assumes all infections are identified for the entire population. Negative values (i.e. increases; see grey cells) are possible due to model stochasticity (N.B., none are significant at $a = 0.05$ level). The black lines delineate cells with a median effectiveness estimate less than 25%.

reinfections will thus underestimate the actual (see the last four panels in figure 2b that discount reinfections, versus those in figure 2a that include reinfections). For the latter, prior epidemic dynamics could affect outcomes measured during a given period and consequently obscure the true effectiveness for masking. As shown in figure 1, epidemic size could vary across seasons, due to e.g. accumulated population immunity. Studies that measure the effectiveness based on a specific time period without accounting for prior dynamics could thus obtain vastly different estimates (figure 2c). These model results demonstrate potential measurement errors and highlight the importance of adopting more comprehensive measurements (e.g. including reinfections) and accounting for prior epidemic dynamics (e.g. examining cumulative effects).

2.5. Additional factors (sensitivity analyses)

The main analysis above assumes that infection risk via non-physical contact is 20% of that via physical contact (i.e. relative risk = 0.2). Depending on the etiologic agent, this relative risk could be higher or lower. In particular, higher infection risk via non-physical contact may apply to pathogens that transmit more efficiently via the airborne route. Thus, we tested the model under different non-physical/physical contact-mediated relative infection risks (0.1, 0.5 and 1, versus 0.2 in the main analysis). Per the POLYMOD contact data, the basic reproduction number R_0 is slightly higher for a relative risk of 1 (3.2 versus 2.5–2.6 for other settings). As expected, it is more challenging to use masking to reduce infection when non-physical contact is as likely as physical contact to cause an infection (electronic supplementary material, figure S1, first row same risk versus other rows with lower relative risks). Nonetheless, the general findings noted in the main analysis are consistent with all other relative risk settings tested (electronic supplementary material, figure S1).

In addition, we hypothesize that mask wearers may be more risk adverse and thus may adopt additional precautionary measures such as reducing contact with others. We tested this by reducing the contact rate of mask wearers by 25%. As expected, the model estimates a further reduction in infections when mask wearers simultaneously reduce their contact rates (electronic supplementary material, figure S2 versus figure 2). For instance, the model estimates a 29% (versus 23% in the main setting) reduction in total infections over 5 years, if half of the population use masks when prevalence is greater than 0.01.

Lastly, over the course of the COVID-19 pandemic, viral mutations have led to more infectious SARS-CoV-2 variants. For instance, the Omicron BA.1 variant, which emerged in the end of 2021, had an R_0 around twice that of the ancestral SARS-CoV-2 variant [16]. To examine the effectiveness of masking against more infectious pathogens, we tested the model under an R_0 of 5.2 (i.e. $2 \times$ the value used in the main analysis, and similar to estimates for the Omicron BA.1 variant [16]). As shown in electronic supplementary material, figures S3 and S4, estimated effectiveness of masking decreases substantially under the much higher R_0 (e.g. estimated reductions of cumulative number of infections over the 5-year study period were half that estimated in the main analysis; see electronic supplementary material, figure S3a versus figure 2a). Similar to the first sensitivity analysis, these results are expected given the increased challenges to curb large epidemic waves caused by more infectious pathogens. Nonetheless, we note that reductions in infection for simulated populations wearing masks are still evident (electronic supplementary material, figures S3 and S4).
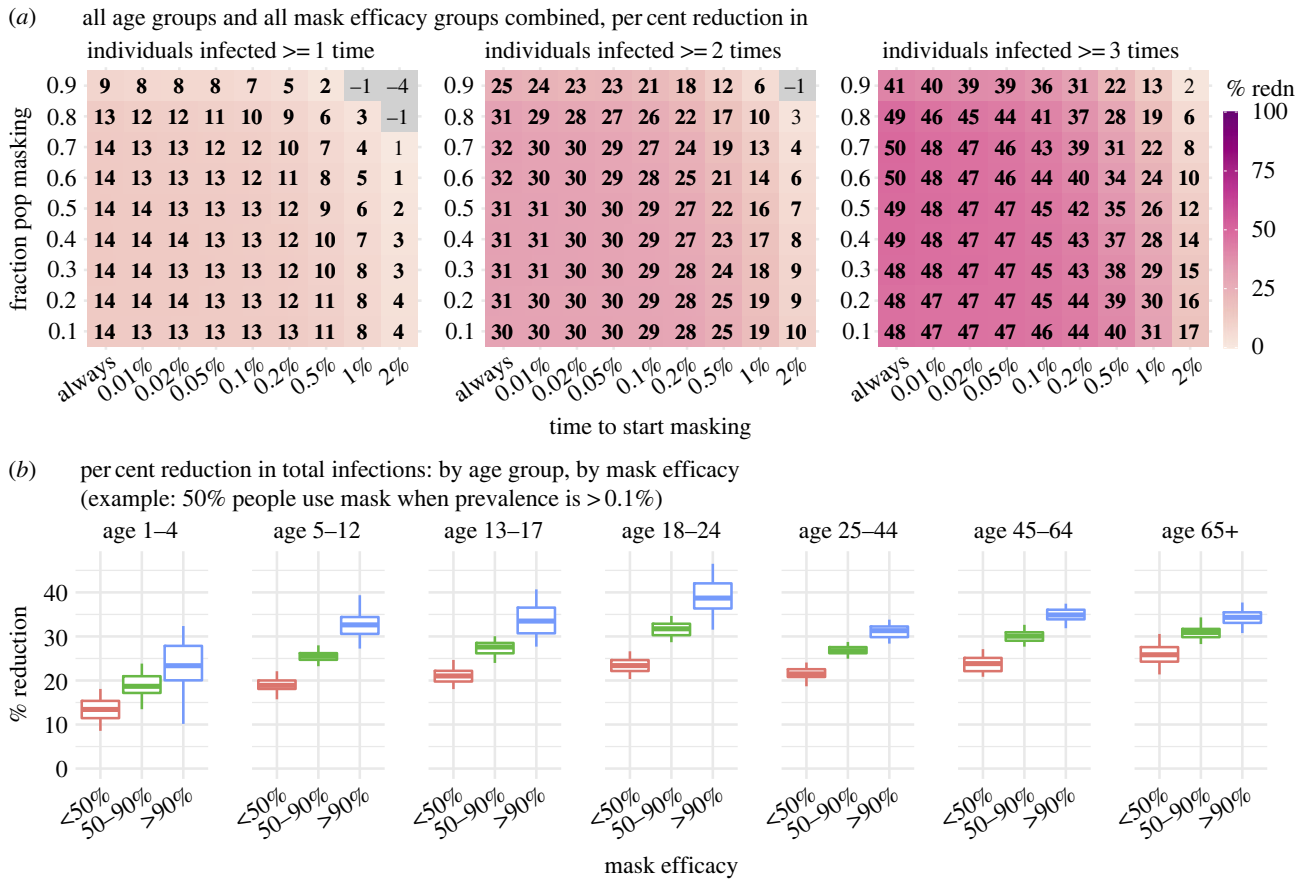
Figure 4. Estimated effectiveness of masking at the individual level (main analysis assuming an R_0 of 2.6 and no masking at home). (*a*) Estimated per cent reduction of repeated infections over the 5-year study period (see panel subtitles), comparing mask wearers with non-wearers in the same simulated population. Numbers in the cells show the median effectiveness estimates (computed per equation (4.3)) for the corresponding masking rate (*y*-axis) and timing of masking (*x*-axis, percentages show population infection prevalence when masking is enacted). Bold fonts indicate the 2.5th% of the estimate is above 0 (i.e. significant at $a = 0.05$ level); note the 95% intervals here are probably narrower than measures in RCTs as the model assumes all infections are identified for the entire population. Negative values (i.e. increases; see grey cells) are possible due to model stochasticity (N.B., none are significant at $a = 0.05$ level). (*b*) Estimated per cent reduction of total infections over the 5-year study period, by age group and mask efficacy. Boxplots show estimated median (middle bar), interquartile range (box edges) and 95% intervals (whiskers).

2.6. Estimated effectiveness of masking at population and individual levels

To examine the effectiveness of masking at the population level, we compare simulated total numbers of infections in a population with different masking settings, with one with no masking. As shown in figure 2*a* and electronic supplementary material, table S2, at the population level, the effectiveness of masking depends on both the masking rate and timeliness of implementation. When implemented early during the epidemic, masking can substantially reduce infections (e.g. approximately 30% estimated reduction over the 5-year study period for a 70% masking rate in the main analysis assuming an R_0 of 2.6), despite the 'leakages' noted above. Model results using parameters based on COVID-19 also show similar reductions in hospitalizations (electronic supplementary material, figure S5 and table S2), suggesting similar downstream impacts on severe disease outcomes. However, timely implementation is crucial and there appears to be a threshold, after which estimated effectiveness declines rapidly—e.g. when community prevalence is greater than 0.2% for the COVID-19-like epidemics tested here (electronic supplementary material, table S2).

To examine the impacts of masking at the individual level, for each simulated population (under a specific masking setting), we compare the (re)infection rates among mask wearers with non-wearers. Across a wide range of settings tested, the model estimates substantial lower risk of reinfection for mask wearers (figure 4*a*). For instance, the model estimates an approximately 50% lower risk of having greater than or equal to 3 infections for mask wearers over the 5-year study period, in the main analysis assuming an R_0 of 2.6. In addition, stratifying by mask efficacy, the estimated protection against infection increases with mask efficacy, and this dose–response is evident across all age groups (figure 4*b*) despite different contact patterns (figure 3*a*).

3. Discussion

Despite the high efficacy of masks in reducing particle inhalation, reported effectiveness of masking has been less than ideal. This large discrepancy has puzzled researchers and prompted doubt on the validity of masking as a public health measure. Here, our model not only generates consistent results reconciling the efficacy-effectiveness discrepancy, but also helps elucidate the underlying drivers. Model results identify factors that can lower effectiveness and highlight a key 'leakage' source via household

transmission. Further, we show the interaction of nonlinear infectious disease transmission dynamics, control measures (e.g. masking and contact reduction/social distancing), and resulting potential measurement biases. As such, the model findings here help provide a more complete picture of masking at both the individual and population levels.

For individuals, modelling results here clearly demonstrate that consistent masking can reduce the risk of initial infection and, over time, reduce the frequency of repeated infection and hence the number of reinfections. For SARS-CoV-2/COVID-19, there remains large uncertainty in the health impact of reinfections. While population-level data have suggested a likely lower severity for reinfections due to prior immune protection (e.g. data from South Africa [17]), individual (re)infection severity may vary vastly. For some individuals, there may be a substantial risk of long-lasting post-acute sequelae following SARS-CoV-2 infection (i.e. long COVID [18]) and the risk of long COVID may further increase for subsequent reinfections. In the light of the unknown long-term health impact of repeated infections, continued precautionary measures (e.g. masking) that reduce repeated infections may be advisable given these potential health impacts.

At the population level, model results here also demonstrate that masking can provide direct protection to mask wearers as well as indirect protection to non-wearers, collectively reducing epidemic intensity. In addition, as for other intervention measures, timely implementation of masking is crucial to ensure its effectiveness. However, consistent with the RCTs, model results here also indicate the overall effectiveness is probably often lower than the reported efficacy measured in laboratory settings. As demonstrated in our counterfactual modelling and testing of a wide range of masking settings, the ultimate effectiveness is limited by multiple factors (particularly, household transmission). It is thus important to recognize this less-than-ideal effectiveness, in order to more effectively devise public health measures (e.g. planning for additional preventive measures).

To focus on the general impact of masking and influencing factors, here we sidestepped several processes shaping SARS-CoV-2/COVID-19 dynamics; these include the emergence and circulation of different variants throughout the pandemic, differential protection of vaccines against different variants, and potential changes in disease severity for subsequent repeated infections. In particular, for more infectious variants (higher R_0), it may be more challenging to reduce transmission via masking, as preliminarily shown in our sensitivity analyses (electronic supplementary material, figures S1 and S3). However, the model developed here provides a general framework to study these questions and beyond.

Despite these limitations, the quantitative modelling developed here helps tease apart the various factors affecting masking effectiveness in the real world. Overall, our results suggest it is important for individuals to use masks during an epidemic, and for policymakers to recognize the non-ideal effectiveness of masking when devising public health measures.

4. Methods

4.1. Data sources and processing

Here, we used contact survey data from the POLYMOD study [15,19] to parametrize individual contact patterns (figure 2a), and mobility data (figure 1c) and vaccination data (electronic supplementary material, figure S6) from New York City (NYC) to model social distancing and vaccination during the pandemic. The POLYMOD study recorded contact data for 7290 participants (97 904 contacts) in eight European countries (e.g. for each contact, household versus non-household, physical versus non-physical and duration) [15,19]. To compute the individual contact duration, we pooled data of all participants, weighted non-physical contact (multiplicatively by a factor 0.2 in the main analysis and 0.1, 0.5 or 1 in the sensitivity analysis) to account for likely lower per-contact infection risk than physical contact. We then summed the weighted contact durations by contact type (household versus non-household) for each individual. Note that an individual's total contact duration per day could exceed one day because simultaneous contacts are counted separately and additively combined. Given reported differences by age group, we further stratified by age group and fit the age-grouped, type-specific contact durations to a probability distribution to formulate the agent-based model. Briefly, for each age group, we fitted the data to six distributions (i.e. uniform, exponential, normal, lognormal, Weibull or gamma) and selected the best-fit distribution based on the Bayesian information criterion [20]. The specific distributions used are shown in electronic supplementary material, table S3.

During the COVID-19 pandemic, many places implemented social distancing policies to reduce population contact rates. To represent these dynamics, as in our previous work [16,21,22], we used mobility data derived from Google Community Mobility Reports [23] to model contact reductions. Here, we used data for NYC during March 2020–June 2021 to formulate the model. From July 2021 onward, we assumed there was no social distancing and increased the mobility linearly to the pre-COVID level over a three-month period and remained at that level for the remainder of the study period (figure 1c).

In addition, vaccines became available in late December 2020 and mass vaccination was rolled out subsequently. To represent the impact of vaccinations, we used data from NYC (four doses, 14 December 2020–9 March 2023, at the time of this study; electronic supplementary material, figure S6) [24] to formulate the model. The first and second doses were administered three to four weeks apart as the primary series and the third and fourth doses were administered more than 6–12 months later. As there is probably little overlap of immune protection from the preceding vaccination series/dose, we additively combined vaccinations of the primary series, third and fourth doses to formulate the daily vaccination rate in the model. Assuming similar vaccination uptakes in the future, we used the mean for the same day of the year over 1 March 2022–9 March 2023 for those under 25 (roughly a 1-year period) and 1 September 2021–9 March 2023 for those aged 25 years and over (longer period given the earlier vaccination rollout for older ages) to formulate vaccination rates for days without data. For simplicity, we assumed a 90% vaccination effectiveness for all doses.

4.2. The agent-based model

The agent-based model simulates individual infection and disease history, depending on age, contact pattern, individual preventive measures (e.g. masking behaviour), prior infection and immune protection, vaccinations and immune protection, and immunity waning. In addition, at the population level, the model incorporates infection seasonality, mobility data to represent social distancing and reductions in contact during the pandemic, and vaccination data to represent mass vaccination during the study period. We describe each component below.

(1) Risk of infection: each day, infection risk for a susceptible individual is computed as the complement of the probability of avoiding infection by household and non-household contacts, i.e.

$$\text{risk of infection} = 1 - \exp\{-[p_0 d_{HH} i_{HH} + p_0 (1 - 1_{\text{mask}} e_{\text{mask}}) d_{nonHH} i_{nonHH} \Delta_c] s\}, \tag{4.1}$$

where p_0 is the per-contact infection probability (set to 50% here); d_{HH} and d_{nonHH} are the total duration of household and non-household contact, respectively, formulated per the POLYMOD data, as described above. The infection prevalence within (i_{HH}) and outside (i_{nonHH}) household are computed by the model; here, for simplicity, we set i_{nonHH} to the population prevalence and i_{HH} to the mask-group-specific prevalence (i.e. assuming similar infection risk for households with similar masking behaviour). For non-household contact, the term $(1 - 1_{\text{mask}} e_{\text{mask}})$ represents the reduction in infection risk due to masking, with 1_{mask} indicating whether the individual wears a mask outside (1 if yes and 0 otherwise) and e_{mask} being the mask efficacy (electronic supplementary material, table S1); Δ_c represents the further reduction through social distancing and was formulated using mobility data as described above (figure 1c). The infection risk may further depend on one's susceptibility (s), e.g. multiplicatively in equation (4.1); here, for simplicity, we set s to 1 for all ages.

(2) Progression of infection: infected (i.e. exposed but not yet infectious) individuals progress through the latency and infectious period exponentially, per: $1 - \exp(-1/D \cdot d)$, where D is the mean sojourn time (i.e. latency or infectious period; electronic supplementary material, table S1), and d the time since the corresponding transition.

(3) Immune protection and waning: due to limited data on immune protection, here we assumed immune protection wanes per a logistic function: $1/(1 + \exp(-k(\tau - 0.5T)))$, where T is the immunity period and τ the time since the corresponding event (i.e. infection or vaccination). To estimate the parameters (k and T), we fit the function to vaccination effectiveness duration data from the UK Health Security Agency (UKHSA) [25] and scaled the estimates to represent the duration of protection against infection (here, $T = 321$ days for prior infection and 180 days for vaccination) and hospitalization ($T = 321$ days for prior infection and 385 days for vaccination). That is, we assumed shorter protection against infection but longer protection against severe disease from vaccination, and moderate protection against infection and severe disease from prior infection.

(4) Vaccination: using the daily vaccination rate computed above (electronic supplementary material, figure S6), we sampled individuals for vaccination according to their infection status (excluding those actively infectious) and vaccination probability based on their time since last vaccination and the estimated immunity period.

(5) Hospitalization: for infected individuals, we computed the probability of hospitalization, based on the age-specific hospitalization rate (h_a; electronic supplementary material, table S1) and immune protection from vaccination ($p_{\text{vac.immunity}}$) and/or prior infection ($p_{\text{inf.immunity}}$), per: $h_a (1 - 1_{\text{vaccinee}} \cdot p_{\text{vac.immunity}})(1 - 1_{\text{recoveree}} \cdot p_{\text{inf.immunity}})$. The level of immune protection $p_{\text{vac.immunity}}$ and $p_{\text{inf.immunity}}$ are computed using the aforementioned logistic function and corresponding immunity period estimates.

(6) Seasonality: to represent the more conducive transmission conditions during winter months for respiratory viruses, including SARS-CoV-2 [26]), we scaled the risk of infection (equation (4.1)) by the seasonal risk (b_s). For simplicity, here we modelled seasonality using a sinusoidal function: $b_s = 1 + a_s \cos(2\pi t/365)$, where a_s is the seasonality amplitude (set to 0.25 here) and t is the day of the year.

(7) Demographics: the model simulates births and deaths (both with a rate set to 1/85 years to maintain the population size) by sampling and replacing those greater than 65 years with newborns (age = 0), and simulates ageing by increasing the age of all individuals by 1 at the end of each calendar year.

4.3. Simulation settings and estimation of masking effectiveness

For all simulations, we initialized the model using 100 000 people and 100 initial infections on 1 March 2020, and ran the model for approximately 5 years (from 1 March 2020 to 31 December 2024); we repeated each simulation 100 times to account for model stochasticity and computed summary statistics [median, interquartile range, and 95% interval (i.e. 2.5% and 97.5% percentile)]. For the main analysis (masking is applied only outside the home; $R_0 = 2.6$), to test the effectiveness of masking under different conditions, we varied the masking rate from 10% to 100% with a 10% increment and assigned masking choice for each individual (all those aged 1 year and over; i.e. no masking for those under 1) using random sampling. For each masking rate, we further varied the timing of masking, from masking outside the home at all times to only when prevalence is above 0.01%, 0.02%, 0.05%, 0.1%, 0.2%, 0.5%, 1% and 2%, separately. In total, there were 90 (10 masking rates × 9 timings) combinations. For comparison, we also simulated a baseline setting with no masking.

For the counterfactual analysis (masking is applied both within and outside the home; $R_0 = 2.6$), we computed the risk of infection as $1 - \exp[-p_0(1 - 1_{\text{mask}} e_{\text{mask}})(d_{HH} i_{HH} + d_{nonHH} i_{nonHH} \Delta_c) s]$, instead of using equation (4.1). For the first sensitivity analysis, we used different weights to scale the relative per-contact infection risk for non-physical contact (0.1, 0.5 or 1, versus 0.2 in the main analysis; $R_0 = 3.2$ when the relative infection risk was set to 1 and 2.5–2.6 for other settings). In the second sensitivity analysis (using the same base R_0 of 2.6), we tested a 25% additional reduction in contact rate for mask wearers, by setting Δ_c in equation (4.1) to $0.75\Delta_c$ for individuals using masks. In the third sensitivity analysis, we set R_0 to 5.2 (i.e. 2× the value used in the main analysis) by assuming an infection probability of 100% (20%) for physical (non-physical) contact (versus 50% and 10% in the main analysis for the two corresponding infection probabilities), and kept all other parameters the same as used in the main analysis.

At the population level, the effectiveness of masking is estimated per the reduction in a given health outcome (infection, repeated infection, hospitalization or repeated hospitalization) under a given masking setting (x_{masking}) compared with the baseline setting without masking (x_0), i.e.

$$E_{\text{pop}} = \frac{x_{\text{masking}} - x_0}{x_0} \times 100\%. \tag{4.2}$$

At the individual (group) level, we compare the rates of a given health outcome (r; i.e. normalized by the number of individuals in each group) among mask wearers versus non-wearers in the same simulated population, i.e.

$$E_{\text{ind}} = \frac{r_{\text{wearers}} - r_{\text{nonwearers}}}{r_{\text{nonwearers}}} \times 100\%. \tag{4.3}$$

All inference, forecast and statistical analyses were carried out using the R language (https://www.r-project.org).

Ethics. This work did not require ethical approval from a human subject or animal welfare committee.

References

1. Lindsley WG, Blachere FM, Law BF, Beezhold DH, Noti JD. 2021 Efficacy of face masks, neck gaiters and face shields for reducing the expulsion of simulated cough-generated aerosols. *Aerosol. Sci. Technol.* **55**, 449–457. (doi:10.1080/02786826.2020.1862409)

2. Pan J, Harb C, Leng WN, Marr LC. 2021 Inward and outward effectiveness of cloth masks, a surgical mask, and a face shield. *Aerosol. Sci. Tech.* **55**, 718–733. (doi:10.1080/02786826.2021.1890687)

3. Ueki H, Furusawa Y, Iwatsuki-Horimoto K, Imai M, Kabata H, Nishimura H, Kawaoka Y. 2020 Effectiveness of face masks in preventing airborne transmission of SARS-CoV-2. *mSphere* **5**, e00637. (doi:10.1128/mSphere.00637-20)

4. Li H *et al.* 2022 Efficacy and practice of facemask use in general population: a systematic review and meta-analysis. *Transl. Psychiatry* **12**, 49. (doi:10.1038/s41398-022-01814-3)

5. Cowling BJ, Escandón JX. 2021 Re: Effectiveness of public health measures in reducing the incidence of COVID-19, SARS-CoV-2 transmission, and COVID-19 mortality: systematic review and meta-analysis. *BMJ* **375**, e068302.

6. Jefferson T, Foxlee R, Del Mar C, Dooley L, Ferroni E, Hewak B, Prabhala A, Nair S, Rivetti A. 2023 Physical interventions to interrupt or reduce the spread of respiratory viruses. *Cochrane Database Syst. Rev.* **2023**, CD006207. (doi:10.1002/14651858.CD006207.pub6)

7. Freyer FJ. 2023 *'It is very hard to study this topic': how research sparked questions around masking and COVID — again*. See https://www.bostonglobe.com/2023/03/02/metro/mask-debate-resumes-do-they-work/.

8. GAVI (The Vaccine Alliance). 2023 *Yes, masks reduce the risk of spreading COVID, despite a review saying they don't*. See https://www.gavi.org/vaccineswork/yes-masks-reduce-risk-spreading-covid-despite-review-saying-they-dont?gclid=CjwKCAjw0ZiiBhBKEiwA4PT9z3uZkCyAqgbVQePamR1Jb0B5geeKqkoWpBdQvTY1DOqzDCzgOldZ6hoCY-wQAvD_BwE.

9. Leech G *et al.* 2022 Mask wearing in community settings reduces SARS-CoV-2 transmission. *Proc. Natl Acad. Sci. USA* **119**, e2119266119. (doi:10.1073/pnas.2119266119)

10. Cheng Y, Ma N, Witt C, Rapp S, Wild PS, Andreae MO, Poschl U, Su H. 2021 Face masks effectively limit the probability of SARS-CoV-2 transmission. *Science* **372**, 1439–1443. (doi:10.1126/science.abg6296)

11. Li Q *et al.* 2020 Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. *N. Engl. J. Med.* **382**, 1199–1207. (doi:10.1056/NEJMoa2001316)

12. Wu JT, Leung K, Leung GM. 2020 Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study. *Lancet* **395**, 689–697. (doi:10.1016/S0140-6736(20)30260-9)

13. Li R, Pei S, Chen B, Song Y, Zhang T, Yang W, Shaman J. 2020 Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV-2). *Science* **368**, 489–493. (doi:10.1126/science.abb3221)

14. Yang W *et al.* 2021 Estimating the infection-fatality risk of SARS-CoV-2 in New York City during the spring 2020 pandemic wave: a model-based analysis. *Lancet Infect. Dis.* **21**, 203–212. (doi:10.1016/S1473-3099(20)30769-6)

15. Mossong J *et al.* 2008 Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLoS Med.* **5**, e74. (doi:10.1371/journal.pmed.0050074)

16. Yang W, Shaman JL. 2022 COVID-19 pandemic dynamics in South Africa and epidemiological characteristics of three variants of concern (Beta, Delta, and Omicron). *Elife* **11**, e78933. (doi:10.7554/eLife.78933)

17. Wolter N *et al.* 2022 Early assessment of the clinical severity of the SARS-CoV-2 omicron variant in South Africa: a data linkage study. *The Lancet* **399**, 437–446. (doi:10.1016/S0140-6736(22)00017-4)

18. Davis HE, McCorkell L, Vogel JM, Topol EJ. 2023 Long COVID: major findings, mechanisms and recommendations. *Nat. Rev. Microbiol.* **21**, 133–146. (doi:10.1038/s41579-022-00846-2))

19. Mossong J *et al.* 2020 *POLYMOD social contact data*. See https://zenodo.org/record/3874557.

20. Schwarz G. 1978 Estimating the dimension of a model. *Ann. Statist.* **6**, 461–464. (doi:10.1214/aos/1176344136)

21. Yang W, Shaman J. 2021 Development of a model-inference system for estimating epidemiological characteristics of SARS-CoV-2 variants of concern. *Nat. Commun.* **12**, 5573. (doi:10.1038/s41467-021-25913-9)

22. Yang W, Shaman J. 2022 COVID-19 pandemic dynamics in India, the SARS-CoV-2 Delta variant and implications for vaccination. *J. R. Soc. Interface* **19**, 20210900. (doi:10.1098/rsif.2021.0900)

23. Google Inc. 2020 *Community Mobility Reports*. See https://www.google.com/covid19/mobility/.

24. New York City Department of Health and Mental Hygiene. 2023 *COVID-19 Vaccination Reporting*. See https://github.com/nychealth/covid-vaccine-data.

25. UK Heath Security Agency. 2022 *COVID-19 vaccine surveillance report (Week 17, 28 April 2022)*. See https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1072064/Vaccine-surveillance-report-week-17.pdf.

26. Wiemken TL, Khan F, Puzniak L, Yang W, Simmering J, Polgreen P, Nguyen JL, Jodar L, McLaughlin JM. 2023 Seasonal trends in COVID-19 cases, hospitalizations, and mortality in the United States and Europe. *Sci. Rep.* **13**, 3886. (doi:10.1038/s41598-023-31057-1)

27. Yang W. 2024 Code for: wan-yang/masking_effectiveness: masking_effectiveness (masking_effectiveness). Zenodo. (doi:10.5281/zenodo.10626876)

28. Yang W, Shaman J. 2024 Reconciling the efficacy and effectiveness of masking on epidemic outcomes. Figshare. (doi:10.6084/m9.figshare.c.7090098)