

## TITLE PAGE

**Title of Manuscript:** Epidemiological impact and cost-effectiveness analysis of COVID-19 vaccination in Kenya

### Authors' information:

Stacey Orangi<sup>1,2\*</sup>, John Ojal<sup>3,4\*</sup>, Samuel P. C. Brand<sup>5,6</sup>, Cameline Orlando<sup>3</sup>, Angela Kairu<sup>1</sup>, Rabia Aziza<sup>5,6</sup>, Morris Ogero<sup>3</sup>, Ambrose Agweyu<sup>3,8</sup>, George M Warimwe<sup>3,8</sup>, Sophie Uyoga<sup>3</sup>, Edward Otieno<sup>3</sup>, Lynette I Ochola-Oyier<sup>3</sup>, Charles N Agoti<sup>3</sup>, Kadondi Kasera<sup>9</sup>, Patrick Amoth<sup>9</sup>, Mercy Mwangangi<sup>9</sup>, Rashid Aman<sup>9</sup>, Wangari Ng'ang'a<sup>10</sup>, Ifedayo M O Adetifa<sup>3,4</sup>, J Anthony G Scott<sup>3,4</sup>, Philip Bejon<sup>3,8</sup>, Matt. J. Keeling<sup>5,6,11</sup>, Stefan Flasche<sup>4</sup>, D. James. Nokes<sup>3,5,6\*\*</sup>, Edwine Barasa<sup>1,2,8\*\*</sup>

\*Stacey Orangi and John Ojal contributed equally.

\*\* D. James Nokes and Edwine Barasa contributed equally

### Affiliations:

<sup>1</sup> Health Economics Research Unit, KEMRI-Wellcome Trust Research Programme, Nairobi, Kenya

<sup>2</sup> Institute of Healthcare Management, Strathmore University, Nairobi, Kenya

<sup>3</sup>Kenya Medical Research Institute (KEMRI) -Wellcome Trust Research Programme (KWTRP), Kilifi, Kenya

<sup>4</sup>The Centre for Mathematical Modelling of Infectious Diseases, London School of Hygiene and Tropical Medicine, London, United Kingdom

<sup>5</sup>The Zeeman Institute for Systems Biology and Infectious Disease Epidemiology Research (SBIDER), University of Warwick, UK.

<sup>6</sup>School of Life Sciences, University of Warwick, UK.

<sup>7</sup>Population Health Unit, Kenya Medical Research Institute -Wellcome Trust Research Programme, Nairobi, Kenya;

<sup>8</sup> Centre for Tropical Medicine and Global Health, Nuffield Department of Medicine, University of Oxford, Oxford, United Kingdom

<sup>9</sup>Ministry of Health, Government of Kenya, Nairobi, Kenya

<sup>10</sup>Presidential Policy & Strategy Unit, The Presidency, Government of Kenya

<sup>11</sup>Mathematics Institute, University of Warwick, UK

\*Correspondence to:

Stacey Orangi. Email: [sorangi@kemri-wellcome.org](mailto:sorangi@kemri-wellcome.org). P.O Box 43640-00100, Nairobi Kenya.

John Ojal. Email: [JOjal@kemri-wellcome.org](mailto:JOjal@kemri-wellcome.org). P.O Box 230-80108, Kilifi, Kenya.

**Keywords:** COVID-19; Kenya; cost-effective analysis; COVID-19 vaccines modelling

## ABSTRACT

**Background:** Few studies have assessed the benefits of COVID-19 vaccines in settings where most of the population had been exposed to SARS-CoV-2 infection.

**Methods:** We conducted a cost-effectiveness analysis of COVID-19 vaccine in Kenya from a societal perspective over a 1.5-year time frame. An age-structured transmission model assumed at least 80% of the population to have prior natural immunity when an immune escape variant was introduced. We examine the effect of slow (18 months) or rapid (6 months) vaccine roll-out with vaccine coverage of 30%, 50% or 70% of the adult (> 18 years) population prioritizing roll-out in over 50-year olds (80% uptake in all scenarios). Cost data were obtained from primary analyses. We assumed vaccine procurement at \$7 per dose and vaccine delivery costs of \$3.90-\$6.11 per dose. The cost-effectiveness threshold was USD 919.

**Findings:** Slow roll-out at 30% coverage largely targets over 50-year-olds and resulted in 54% fewer deaths (8,132(7,914 to 8,373)) than no vaccination and was cost-saving (ICER=US\$-1,343 (-1,345 to -1,341) per DALY averted). Increasing coverage to 50% and 70%, further reduced deaths by 12% (810 (757 to 872) and 5% (282 (251 to 317) but was not cost-effective, using Kenya's cost-effectiveness threshold (\$ 919.11). Rapid roll-out with 30% coverage averted 63% more deaths and was more cost-saving (ICER=\$-1,607 (-1,609 to -1,604) per DALY averted) compared to slow roll-out at the same coverage level, but 50% and 70% coverage scenarios were not cost-effective.

**Interpretation:** With prior exposure partially protecting much of the Kenyan population, vaccination of young adults may no longer be cost-effective.

## KEY QUESTIONS

### What is already known?

- The COVID-19 pandemic has led to a substantial number of cases and deaths in low-and middle-income countries.
- COVID-19 vaccines are considered the main strategy of curtailing the pandemic. However, many African nations are still at the early phase of vaccination.
- Evidence on the cost-effectiveness of COVID-19 vaccines are useful in estimating value for money and illustrate opportunity costs. However, there is a need to balance these economic outcomes against the potential impact of vaccination.

### What are the new findings?

- In Kenya, a targeted vaccination strategy that prioritizes those of an older age and is deployed at a rapid rollout speed achieves greater marginal health impacts and is better value for money.
- Given the existing high-level population protection to COVID-19 due to prior exposure, vaccination of younger adults is less cost-effective in Kenya.

### What do the new findings imply?

- Rapid deployment of vaccines during a pandemic averts more cases, hospitalisations, and deaths and is more cost-effective.
- Against a context of constrained fiscal space for health, it is likely more prudent for Kenya to target those at severe risk of disease and possibly other vulnerable populations rather than to the whole population.

## INTRODUCTION

As of early April 2022, Kenya has experienced five distinct waves of the COVID-19 pandemic with more than 320,000 reported cases and 5,600 deaths.[1] While at the global level vaccines to prevent severe disease from SARS-CoV-2 are the main strategy for curtailing the pandemic burden on health,[2] most African nations are still at a very early phase of vaccine roll-out, particularly in tropical sub-Saharan Africa, with most countries at less than 10% of the adult population fully vaccinated.[1] However, in contrast to other part of the world where low vaccine coverage in high risk groups has led to high mortality even from the omicron variant,[3] in Kenya cross-sectional serological surveys of anti-SARS-CoV-2 spike protein antibodies together with transmission dynamic model forecasts indicate that about 80% of the population have been exposed to the virus at least once and thus generated considerable immunity[4] with similar estimates in the region.[5] This raises the question what additional benefit can vaccination still have in mitigating future disease burden from COVID-19?

The Kenyan government is pursuing a phased COVID-19 vaccination strategy that aims to follow a risk-prioritization matrix leading sequentially to the vaccination of all adults by December 2022.[6] The prioritized population are an estimated 30% of the adult population and include health and other essential workers, individuals at high risk of severe disease (those above 58 years, and those above 18 years with co-morbidities), and individuals at high risk of infection (individuals in congregate settings, and those working in hospitality and transport sectors).[6] Vaccine roll-out commenced in early March 2021. As of early April, 2022 more than 17.7 million doses had been administered with 30% of Kenya's adult population above the age of 18 years being fully vaccinated.[7] The initial procurement comprised of the Oxford/Astra Zeneca vaccine mainly sourced through the COVID-19 Vaccines Global Access Facility (COVAX) mechanism and bilateral negotiations, evolving more recently to a multi-vaccine type deployment through additional sources including the African Union's (AU) African Vaccine Acquisition Task Team (AVATT) mechanism.[6]

Economic evaluations are useful in providing evidence of the value for money for different health interventions and illustrates the opportunity costs of the interventions in a setting with many competing priorities. However, there is a need to balance these economic outcomes against the potential impact of the interventions. Therefore, this study evaluates the potential epidemiological impact and cost-effectiveness of different vaccine roll-out scenarios in a Kenyan population that has already acquired a high-level immunity due to prior infections. The study employs a partially retrospective perspective with vaccination scenarios beginning September 2021 and with an immune escape variant striking in November 2021.

## METHODS

### Study design

This study is an impact and cost-effectiveness analysis of COVID-19 vaccine roll-out strategies that uses cost estimates from primary costing studies and vaccine effectiveness measures from an age structured transmission model. The costs and effects are estimated from a societal perspective for a period of 1.5 years (1<sup>st</sup> September 2021 to 28<sup>th</sup> February 2023) starting at the peak of the Kenyan delta wave and simulating the emergence of a partial immune escape variant (omicron-like) from November 2021.

## Intervention comparators

### i. Primary analysis

We carry out an incremental analysis of four vaccination coverage scenarios deployed over an 18-month period (non-rapid deployment), starting at 0% coverage in September 2021 (Table 1): No vaccination (0% coverage), or 30%, 50% and 70% coverage of the population older than 18 years with prioritization of those aged 50 years and above (until 80% of those >50y old are fully vaccinated), then the remaining doses given to those 18-49y old).

### ii. Secondary analysis

We consider a secondary analysis that assesses the same scenarios under the primary analysis but with rapid vaccine deployment in which the targeted vaccine coverage is attained within 6 months of starting vaccination.

An assumption was made that all the vaccination coverage scenarios and deployment strategies were implemented alongside a low intensity mix of non-pharmaceutical interventions (NPI). The low intensity NPI is matched with how government progressively lifted or modified the restrictions and refers to re-opening of international borders, relaxed curfew, controlled public gatherings, controlled re-opening of restaurants and bars, controlled re-opening of schools, ban lift on mandatory use of masks, and adherence to hand hygiene from November 2020 to the time of writing this manuscript.

**Table 1:** Intervention comparators and number vaccinated within 1.5 years' time horizon

Vaccination strategy	Number of >50y olds who were vaccinated (proportion of those>50y)	Number of 18-49y olds who were vaccinated (proportion of those 18-49y)
No vaccination		
30% adult coverage strategy	4,133,775 (80%)	3,186,225 (19%)
50% adult coverage strategy	4,133,775 (80%)	8,366,225 (41%)
70% adult coverage strategy	4,133,775 (80%)	13,366,225 (65%)

## Transmission modelling and parameter inference

We extended a dynamic SARS-CoV-2 transmission model previously designed to estimate population level immunity from natural infection in Kenya by fitting to case notification and serological data[3] to include additional age structure and vaccination status. In common with other approaches to modelling SARS-CoV-2 transmission,[8,9] we assume that the rate of new infections depends on: (i) age and setting-specific contact rates within the population, (ii) frequency of Alpha, Beta and Delta variants of SARS-CoV-2 among the infected sub-population, (iii) the first and second dose vaccine protection against infection in each age group which were assumed to wane over time, and, (iv) prior primary infections. The probability of being infected with SARS-CoV-2 per infectious contact, and the chance of developing symptoms upon infection, increased substantially with age (see *supplementary information* for details of the transmission model).

The goal of the transmission model is to project the health gains of the vaccine deployment strategies described above in comparison with the no vaccine scenario. This requires estimation of parameters pertaining to the risk of transmission, and, of risk factors associated with infection given age, and the

infecting variant of SARS-CoV-2 in the Kenyan setting. These parameters were inferred by fitting the model to the following Kenyan epidemiological data (see SI for inference methods):

- Daily reported numbers of positive and negative PCR tests from the Kenyan Ministry of Health COVID-19 linelist (between 1<sup>st</sup> January 2021 and 1<sup>st</sup> November 2021).
- Cross-sectional serological surveys of (a) donor samples from the Kenyan National Blood and Transfusion Service (KNBTS)[10] and (b) demographic surveillance systems (between 1<sup>st</sup> January 2021 and 27<sup>th</sup> May 2021).[11]

We used a Bayesian hierarchical inference approach aimed at allowing inference on reporting fraction in counties with higher numbers of serological tests to influence inference of reporting fraction in counties with lower numbers of serological tests (see *supplementary information* for details on underlying data for age-specific effects and details on inference methodology).

### **Infection outcome modelling and risk factor inference**

Bayesian inference of transmission model parameters generated a posterior predictive distribution for the number of SARS-CoV-2 infections in Kenya broken down by day, county, age of infected, infecting variant of SARS-CoV-2, and, whether it was a primary infection event or a re-infection event. We categorized the outcome of each infection as being either deadly, critical (requiring treatment in an intensive care unit (ICU)), severe (requiring in-patient hospitalization in a general ward), mild or asymptomatic. Severe and critical infections were assumed to cause admission to a health facility's general ward or ICU for an average of 12 days post-infection. Severe infection was assumed to lead to an average 7 day stay in a general COVID ward before discharge. Critical infection leads to an average 7 day stay in ICU,[12] before transfer to a general COVID ward for a further average 7 day stay before discharge (see *supplementary information* for details on hospital durations of stay).

Risk factors for infection outcome were inferred using reported Kenya outcome data:

- Daily reported numbers occupying general health facilities with COVID-19 as the diagnosed cause (1<sup>st</sup> March 2021 – 1<sup>st</sup> November 2021).
- Daily reported numbers occupying ICUs with COVID-19 as the diagnosed cause (1<sup>st</sup> March 2021 – 1<sup>st</sup> November 2021).
- Daily reported incidence of death with COVID-19 as the diagnosed cause (1<sup>st</sup> January 2021 – 1<sup>st</sup> November 2021).

### **Vaccination rollout modelling**

We used the fitted model to predict the course of the pandemic from 1<sup>st</sup> September 2021 (historically this was past the peak of the fourth wave of cases in Kenya) to 30 June 2023 and the impact of vaccination on, severe and critical disease, and deaths. We distribute the total number of doses planned under each vaccination scenario to the 47 counties proportionally according to population size above the age of 18 years.[13] We assume that the number of doses given per day will be the same during the study period. Doses will be offered to adults older than 50 years first, until take up of available vaccines dropped off, which we assumed would occur once 80% of over 50s had taken up both doses. The remaining doses will subsequently be randomly allocated to all 18- to 50-year-olds. Within the model, individuals are either unvaccinated, partially vaccinated (14 days after receipt of the first dose), fully vaccinated (14 days after receipt of the second dose) or have waned vaccine effectiveness. We assumed vaccine effectiveness against death (delta variant) to range from 90% to 95% after the first dose and 95% to 99% after the second dose.[14] Vaccine effectiveness against severe



or critical disease (delta variant) of 80% to 90% and 95% to 99% after the first and second dose respectively.[14] The vaccine effectiveness against acquisition of infection per infectious contact (delta variant) was 55% to 65% and 65% to 80% after the first and second dose respectively.[14] We assumed an effectiveness of 0% to 35% and 0% to 69% against onward transmission, per infection (delta variant), after either the first or second dose.[15] We assume that immunity due to either past infection of vaccination eventually wanes to 70% protection against disease and 0% protection against infection, with a mean time to complete waning of 460 days after the second dose of vaccine and 5 years following natural infection.[16] Furthermore, we assume that protection due to prior infection combined constructively with vaccination; that is that people who had previously had a natural infection episode of SARS-CoV-2 were further protected from reinfection by vaccination (see *supporting information*).

### **Immune escape variant**

The scenarios investigated in this paper involve the rapid spread of a new variant of SARS-CoV-2 that, due to evolutionary adaptation, partially avoids protection from infection due to prior naturally acquired immunity and/or vaccination. Concretely, we assume that the immune escape variant enters Kenya in early November 2021 and rapidly dominates transmission by 15<sup>th</sup> November 2021. Compared to homologous protection against reinfection with the Delta variant, the protection afforded by prior infection and/or vaccination against acquiring the novel immune escape variant is assumed to be decreased by 50%, with all epidemiological rates increased such that the mean generation time of transmission is reduced by 30% compared to the transmission of the Delta variant. However, we also assume that the fundamental reproductive number and risk factors for severe, critical and deadly outcomes are unchanged compared to Delta. (see *supplementary information* for the details of how a 50% decrease in protection from infection was implemented).

### **Cost estimates**

The cost estimates used in this study were derived using a hybrid method that involved both an ingredients approach (bottom-up) and a top-down approach.[17,18] The analysis used economic costs, which reflect the opportunity cost and incorporated both recurrent and capital costs. Capital costs were annuitized using a discount rate of 3% over their useful life. Costs incurred in other years were adjusted for inflation using the Gross domestic product (GDP) deflator and reported in 2021 United States Dollars (USD). Key model cost input parameters are shown in Table 2 and the three main cost components are described below. The costs of NPIs were excluded as all vaccination strategies employed the same NPI regime (low NPI intensity) and would therefore not change the reported incremental cost-effectiveness ratios (ICERs).

#### **i. Vaccination costs**

We included vaccine and related supplies costs, as well as vaccine delivery costs. Vaccine and related supply costs were the economic costs to purchase the vaccine and related supplies such as syringes and safety boxes through the COVAX facility. A base cost of \$7 was used for vaccine procurement, which is the country's procurement cost from the COVAX facility. Additionally, the freight costs, insurance costs, import declaration fees, clearance fees, and the railway development levy associated with the vaccines and its supplies were included. Vaccine and syringe wastage rates of 10% were assumed.[19] Vaccine delivery costs referred to costs associated with delivering COVID-19 immunizations to the adult population and were estimated across six components 1) vaccine supply chain 2) vaccine safety monitoring and adverse events following immunization management 3)

training 4) advocacy, communication and social mobilization 5) data management, monitoring and supervision 6) vaccine administration. The resource used and costs were estimated through the analysis of programmatic budgets, and through key informant interviews. Details of the vaccine procurement and delivery cost analysis and results are reported elsewhere.[19] This analysis assumed equivalent vaccine delivery costs for both the rapid and non-rapid vaccination strategies.

ii. Treatment costs

The direct medical costs of COVID-19 treatment were sourced from a recently conducted study that examined the unit costs for COVID-19 case management in Kenya.[20] This costing analysis employed an ingredients-based approach to estimate health care costs across the disease severity categories; with the exclusion of adverse events costs.[20]

iii. Productivity losses

Productivity losses due to illness and mortality were estimated using a human capital approach.[21] The impact of COVID-19 on lost time through illness or morbidity was estimated by accounting for the average Kenyan's productivity measure (GDP per capita) and duration of disease/duration of quarantine; the latter was used where duration of illness was less than the 14 day quarantine period in Kenya. For asymptomatic and mild disease, the testing rate was accounted for and an assumption was made that only those in the informal sector are likely not to be productive as they isolate. Further, the economic impact of COVID-19-related mortality was estimated by considering the years of life lost because of premature mortality and the average productivity measure. We did not account for productivity losses from long COVID, as the burden is poorly defined in our setting. (See equation (a) in the supplementary information)

### **Disability-adjusted life-years**

The outcome of the cost-effectiveness analysis was reported in terms of disability-adjusted life-years (DALYs); the sum of years of life lost (YLL) and years lost due to disability (YLD) [22]. (See equations (b), (c), and (d) in the supplementary information).

DALYs were calculated considering a discount rate of 3%, the Kenyan 2019 standardized life expectancies,[23] assumed duration of illness of 7 days for asymptomatic and mild disease and 12 and 20 days for severe and critical disease respectively,[24] as well as disability weights. COVID-19 is a novel disease, and its disability weights are currently not available. Therefore, for asymptomatic COVID-19 disease we assumed a disability weight of 0. For mild-to-moderate COVID-19 symptoms and severe disease, we used disability weights from the 2013 Global Burden of disease of 0.051 (0.032-0.074) and 0.133 (0.088-0.190) assigned to infectious disease with moderate acute episodes and severe episodes respectively.[25] For critical disease, we assume disability weights of 0.655 (0.579-0.727) assigned to intensive care unit admissions.[26] This analysis did not incorporate age-weighting in the DALYs. These input parameters are reported in Table 2.

The incremental cost-effectiveness ratio (ICER) was the measure of cost-effectiveness calculated as the net change in total costs and DALYs averted between comparators. The ICER was compared with the opportunity cost-based on Kenya's cost-effectiveness threshold (USD 919.11).[27,28]

$$ICER = (Cost_{index} - Cost_{baseline}) / (DALY_{baseline} - DALY_{index})$$

Where:



$Cost_{index}$  = cost of strategy of interest

$Cost_{baseline}$  = cost of the next less effective strategy

$DALY_{index}$  = total DALYs under the strategy of interest.

$DALY_{baseline}$  = total DALYs under the next less effective strategy

ICERs are estimated within each of the two roll-out scenarios slow and rapid and are not comparable between the two vaccine-deployment cases except when the baseline is no vaccination.

**Table 2:** Key analysis parameters

Parameter	Values (Lb; Ub)	Source
<b>Cost-effectiveness parameters</b>		
Treatment costs (2021 US\$)		
Per day, per patient unit cost of management of asymptomatic COVID-19	\$19.75*testing rate	[20]
Per day, per patient unit cost of management of mild to moderate COVID-19	\$19.75*testing rate	[20]
Per day, per patient unit cost of management of severe COVID-19	\$129.45	[20]
Per day, per patient unit cost of management of critical COVID-19	\$623.14	[20]
Testing rate		
*Testing rate in the population	0.52%	Proportion of reported to modelled cases
Vaccination costs (2021 US\$)		
Vaccine procurement costs per dose	\$8.67 (Base cost: \$7 and including importation costs)	[19]
Supplies procurement costs per dose	\$0.08	[19]
Vaccine delivery cost per dose (no vaccination)	\$0	[19]
Vaccine delivery cost per dose (30% coverage)	\$6.11	[19]
Vaccine delivery cost per dose (50% coverage)	\$4.16	[19]
Vaccine delivery cost per dose (70% coverage)	\$3.90	[19]
Duration of disease and length of hospitalization		
Length of hospitalization for severe episode	7 days (4-11)	Assumption
Length of ICU stay for critical episode	7 days (4-11)	[12]
Duration of asymptomatic disease	7 days	Assumption
Duration of mild to moderate disease	7 days	Assumption
Duration of severe disease	12 days	[24]
Duration of critical disease	20 days	[24]
DALYs		
Disability weight for asymptomatic episode	0	
Disability weight for mild/moderate episode	0.051 (0.032; 0.074)	[25]
Disability weight for severe episode	0.133 (0.088; 0.191)	[25]
Disability weight for critical episode	0.655 (0.579; 0.727)	[26]
*Average age at death		[13]
0-19 years	9.27years	
20-49 years	31.75 years	

50-59 years	54.10 years	
60-69 years	63.85 years	
70-79 years	73.41 years	
80+ years	86.00 years	
Life expectancy		[29]
0-19 years	64.10 years	
20-49 years	40.94 years	
50-59 years	24.71 years	
60-69 years	17.64 years	
70-79 years	11.41 years	
80+ years	4.84 years	
Cost-effectiveness threshold per DALY averted	0.5*GDP per capita. [GDP per capita = \$1,838.21]	[27,28,30]
<b>Transmission dynamic model parameters</b>		
Transmission dynamic model values	See Table S1	

\*Testing rate: A proxy estimate is used that is calculated as a proportion of reported cases to modelled cases across all severity levels from Jan1 to Sep 19, 2021.

\*Average age at death is based on the weighted mean age across the different age groups

### Sensitivity analysis for the model

Vaccine effectiveness against different epidemiological outcomes such as the acquisition of disease, onward transmission, severe disease and death does vary with age, duration between vaccination and testing of efficacy, variant of infection, and type of vaccine being used amongst other factors.[31-34] Therefore, to determine the robustness of the epidemiological model predictions to the vaccine effectiveness parameter values, we performed a sensitivity analysis across a range of values using a vaccine waning effectiveness model fitted to the UK Health Security Agency (UKHSA) COVID-19 data.[14,15]

A univariate sensitivity analysis was done on the economic model to determine the robustness of the unit cost estimates with variations in vaccine procurement costs (base cost of \$3 and \$10 used) and discounting rates of DALYs (rate of 0% used). Further, given the current evidence gap to confidently determine the magnitude of underreporting of COVID-19 deaths,[35] the baseline cost-effectiveness analysis assumed an under-reporting of hospitalization and deaths by a factor of 5 and a one-way sensitivity analysis was done by varying the under-reporting factor (1-4).

A probabilistic sensitivity analysis to explore the influence of some economic parameters on the ICERs was done using Sobol sampling and was based on the statistical distributions in Table S2. Sobol sequences belong to the family of quasi-random sequences which are designed to generate samples of multiple parameters as uniformly as possible over the multi-dimensional parameter space.[36] For the parameters used in the probabilistic sensitivity analysis, the statistical distributions were chosen to model the available prior knowledge represented by existing data, as reported in Table 2. For the cost estimates range, a 20% increase or decrease was assumed for the parameters.

### Stakeholder engagement

The results of the study have been disseminated to key policy makers and relevant stakeholders involved in COVID-19 vaccine deployment in Kenya.

## RESULTS

### Clinical Impacts of Vaccination Strategies and scenarios

The non-rapid deployment of vaccinating 30% of the adult population results in 10% (32 (24 to 38) per 100,000) fewer infections, 54% (8,132 (7,914 to 8,373) fewer deaths compared to no vaccination, and 978 (949 to 1,005) people would need to be vaccinated to prevent 1 death. An increase of vaccine coverage of the adult population to 50% results in a further 1% (4 (3 to 5) per 100,000) reduction in infections, a further 12% (810 (757 to 872) reduction in deaths, and 5,617 (5,218 to 6,011) more people would need to be vaccinated to prevent an additional death. Similarly, an increase of vaccine coverage to 70% leads to a 1% reduction in cases, a 5% reduction in deaths, and 17,730 (15,773 to 19,920) more people would need to be vaccinated to prevent an additional death compared to the 50% vaccination coverage.

In the rapid vaccine rollout strategy, the 30% vaccine coverage averts 12% of cases preventing an average of 39 (29 to 48) per 100,000 infections and 63% of deaths saving an average of 9,433 (9,197 to 9,711) lives compared to no vaccination. Therefore, 843 (819 to 864) people would need to be vaccinated to prevent a death. The 30% coverage under a rapid deployment averts more cases and saves more lives compared to a non-rapid rollout with the same level of coverage. See Table 3 and Figure 1.

### Cost-Effectiveness of Vaccination Strategies

Table 3 shows the total costs, DALYs and ICERs of the vaccination scenarios considered in the analysis from a societal perspective. Under the non-rapid vaccination scenario, vaccinating 30% of the adult population is cost-saving (ICER=\$-1,343 (-1,345 to -1,341) per DALY averted) and hence highly cost-effective. Increasing vaccine coverage to 50% of the adult population was not cost-effective (ICER=\$3,291 (3,287 to 3,295) per DALY averted) compared to 30% coverage. Similarly, increasing vaccine coverage to 70% was deemed not cost-effective (ICER=\$22,623 (22,602 to 22,645) per DALY averted) compared to 50% coverage at a cost-effectiveness threshold of \$919.11.

Under the rapid vaccination scenario, a 30% vaccine coverage strategy was even more cost-effective ICER=\$-1,607 (-1,609 to -1,604) per DALY averted compared to the same coverage level under the non-rapid scenario, and hence is more cost-effective. The ICERs of 50% and 70% coverage strategies under the rapid scenario are \$18,257 (18,226 to 18,287) and \$44,250 (44,126 to 44,374) per DALY averted compared to 30% and 50% coverage strategies respectively and hence are not cost-effective.

### Sensitivity Analysis

Table 3 presents the univariate sensitivity analysis of under-reporting of hospitalizations and deaths, from a societal perspective. Assuming no under-reporting or adjusting the under-reporting factor to 2, results in all the scenarios having ICERs above the cost-effectiveness threshold, except the 30% coverage with a rapid deployment. On the other hand, with an under-reporting factor of 3 or 4, 30% coverage with a rapid and non-rapid vaccination scenario remained cost-saving.

Figure S6 summarizes the effects of vaccine prices and discounting rates of DALYs on the ICER. Vaccine prices, of the two parameters had the largest effect on the ICERs: leading to a 32-103% decrease and a 36-77% increase in ICERs across the different vaccination scenarios.

The one-way sensitivity analysis focusing solely on a health system's perspective is presented in Table S3. When considering this perspective, the total costs across the vaccination strategies increase as coverage increases, as reported from a societal perspective. However, the no vaccination scenario affords the least costs (\$313 million). The reported ICERs increase with increased coverage and the 30% coverage with a non-rapid and rapid vaccination pace are below the threshold: ICER=\$555 (553 to 557) and \$291 (290 to 295) per DALY averted respectively and considered cost-effective from a health system's perspective.

Figure 2 represent the findings of the probabilistic sensitivity analysis from a societal perspective. The region below the cost-effectiveness threshold line and within the grey region, shows all the points that are cost-effective at a cost effectiveness threshold of \$919.11. For instance, the dominance of the 30% coverage scenarios (i.e. more effectiveness at a lower cost) compared to no vaccination, was shown in 100% of the replications (i.e. 100% of the cost-effect pairs were in the southeast quadrant). Further, 100% of the replications for 50% coverage and 70% coverage strategies (both rapid and non-rapid rollout) were in the northeast quadrant (implying that these strategies were more costly but also more effective compared to the 30% and 50% coverage strategies respectively).

Figure 3 presents the cost-effectiveness acceptability curves (CEAC) of the analysis from a societal perspective based on a range of cost-effectiveness thresholds. Under the non-rapid vaccination rollout and given a \$3,300 willingness to pay threshold, the probability of the 50% coverage strategy being cost-effective compared to 30% coverage would be 0.5. Further, there was 0.5 probability that the 70% coverage in comparison to 50% coverage would be cost-effective at a threshold of \$22,600 in the non-rapid deployment.

**Table 3:** Projected clinical outcomes, costs, and the cost-effectiveness of different vaccination strategies in Kenya from a societal perspective

	Health Outcomes		Economic Outcomes			One-way sensitivity analysis of under-reporting in hospitalizations and deaths (factor of under-reporting)			
	*Averted SARS-CoV-2 infections per 100,000 Median (2.5 <sup>th</sup> to 97.5 <sup>th</sup> percentile)	*Averted SARS-CoV-2 deaths Median (2.5 <sup>th</sup> to 97.5 <sup>th</sup> percentile)	Total costs (\$ millions), Median (2.5 <sup>th</sup> to 97.5 <sup>th</sup> percentile)	Total DALYs (thousands) Median (2.5 <sup>th</sup> to 97.5 <sup>th</sup> percentile)	*ICER, (\$ per DALY averted) Mean (95%CI)	ICER (4)	ICER (3)	ICER (2)	ICER (1)
<b>Non-rapid vaccination strategy (administered within 1.5years)</b>									
No vaccination	-	-	787 (740 to 882)	247 (243 to 252)	-	-	-	-	-
30% coverage	32 (24 to 38)	8,132 (7,914 to 8,373)	614 (589 to 659)	114 (110 to 118)	-1,343 (-1,345 to -1,341) Dominant	-905 (-907 to -902) Dominant	-175 (-178 to -173) Dominant	1,278 (1,275 to 1,280)	5,595 (5592 to 5598)
50% coverage	4 (3 to 5)	810 (757 to 872)	658 (636 to 699)	101 (97 to 104)	3,291 (3,287 to 3,295)	4,908 (4,903 to 4,913)	7,598 (7,592 to 7,605)	12,958 (12,949 to 12,967)	28,878 (28,860 to 28,896)
70% coverage	2 (1 to 3)	282 (251 to 317)	763 (742 to 801)	96 (92 to 100)	22,623 (22,602 to 22,645)	29,075 (29,048 to 29,101)	39,791 (39,756 to 39,827)	61,093 (61,040 to 61,146)	123,967 (123,863 to 124,072)
<b>Rapid vaccination strategy (administered within 6 months)</b>									
No vaccination	-	-	787 (740 to 882)	247 (243 to 252)	-	-	-	-	-
30% coverage	39 (29 to 48)	9,433 (9,197 to 9,711)	545 (524 to 582)	93 (89 to 96)	-1,607 (-1,609 to -1,604) Dominant	-1,230 (-1,232 to -1,227) Dominant	-603 (-605 to -600) Dominant	646 (644 to 649)	4,357 (4,354 to 4,360)
50% coverage	1 (0.5 to 2)	250 (201 to 296)	620 (599 to 655)	88 (85 to 92)	18,257 (18,226 to 18,287)	23,582 (23,543 to 23,620)	32,440 (32,389 to 32,491)	50,096 (50,020 to 50,173)	102,582 (102,430 to 102,733)
70% coverage	0.5 (-0.05 to 1)	161 (106 to 208)	731 (713 to 765)	86 (82 to 89)	44,250 (44,126 to 44,374)	56,074 (55,919 to 56,230)	75,768 (75,560 to 75,976)	115,100 (114,788 to 115,413)	232,667 (232,036 to 233,298)

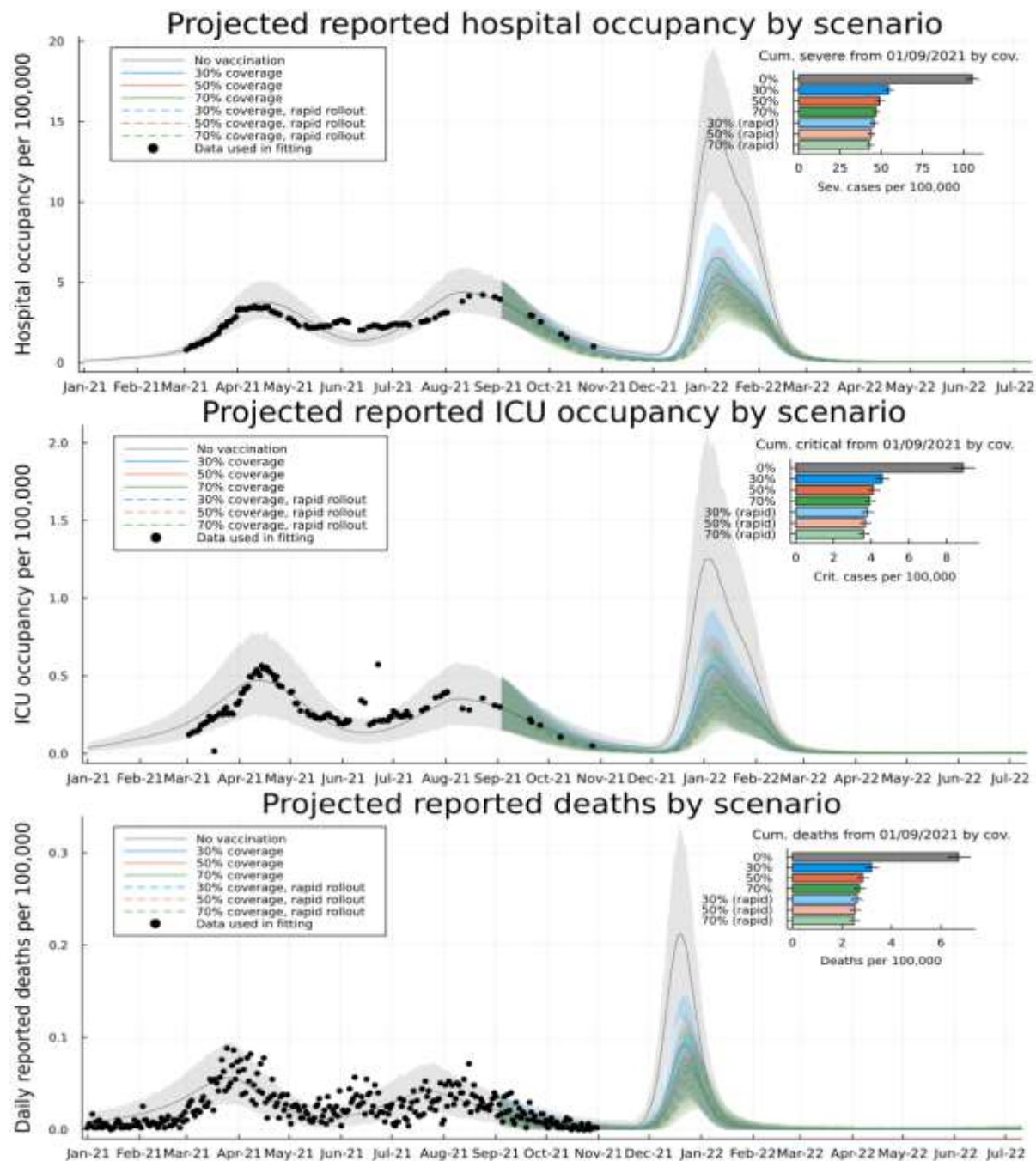
*\*Averted SARS-Cov-2 infections and deaths = This is the incremental averted infections/deaths compared to the vaccination strategy that appears in the row above.*

*\*ICER=Baseline ICER used in analysis where under-reporting in hospitalizations and deaths is adjusted with a factor of 5.*

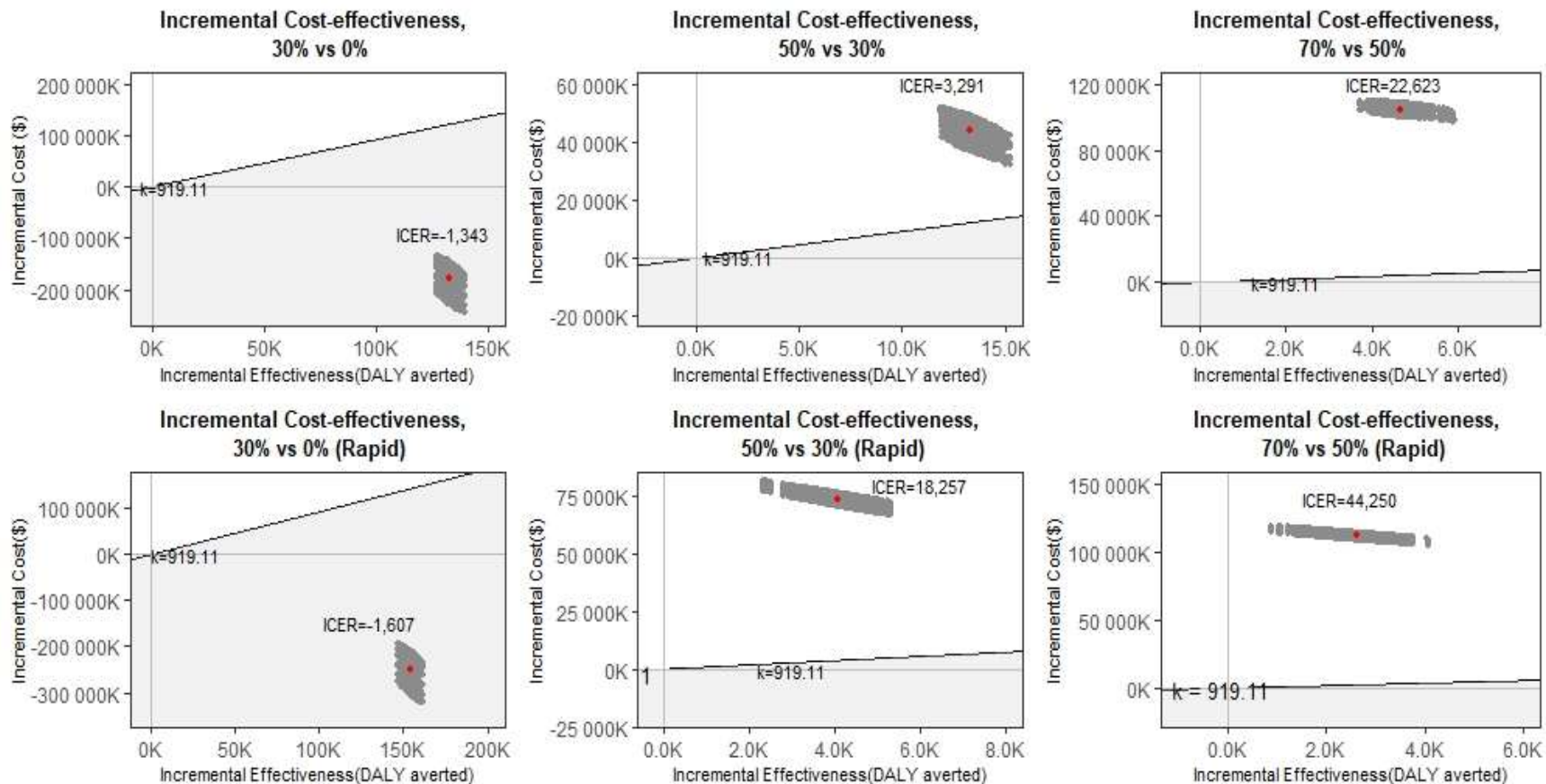
*ICER (4), ICER (3), ICER (2), and ICER (1)= under-reporting factors for hospitalization and deaths used were 4, 3, 2, and 1 respectively.*

*Total averted infections=rounded off to the nearest 100,000; Total Cost=rounded off to the nearest 1,000,000; Total DALY rounded off to the nearest 1,000; Total deaths and ICERs=rounded off to the nearest whole number*

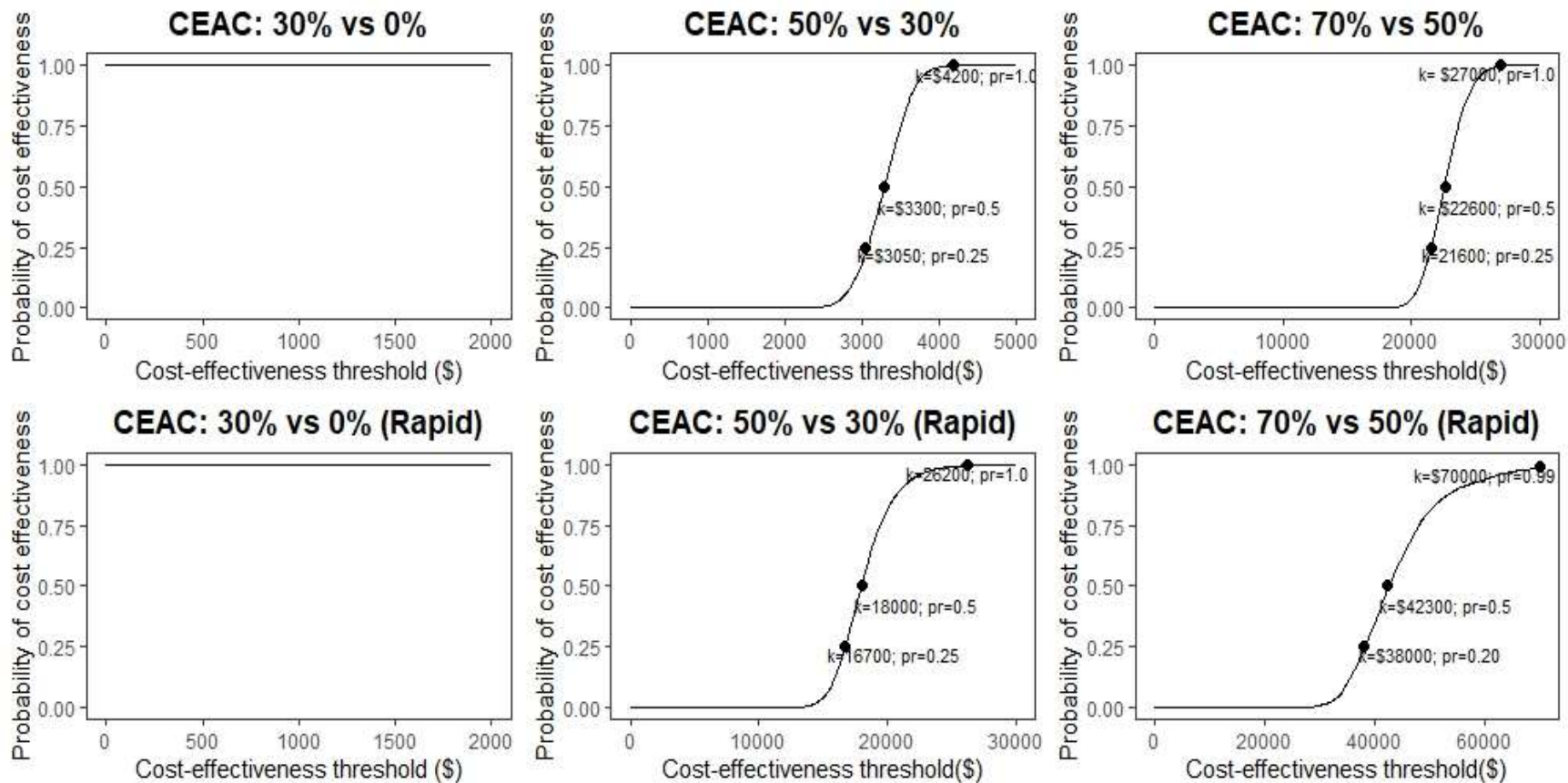




**Figure 1:** Model-based projections and vaccine scenarios: Model-based prediction intervals for daily occupancy of general wards in health facilities in Kenya (top), daily occupancy of intensive care units in Kenya (middle), and daily reported incidence of death with COVID in Kenya (bottom). All scatter points represented data used in inference of the infection outcome model. Grey curves are the posterior mean model prediction (background shading 95% CIs) with no vaccinations. Colored curves represent a target of 30% (blue), 50% (red) and 70% (green) of over 18 year old population in Kenya over 18 months (solid) or 6 months (dashed). Insets: Projections of cumulative number of severe (top), critical (middle) and deadly (bottom) cases after 1<sup>st</sup> September 2021 under each vaccine target scenario



**Figure 2:** Probabilistic sensitivity analysis of different vaccination strategies from a societal perspective. The first row shows the vaccine scenarios comparisons under a non-rapid rollout pace while the second row shows the rapid roll out results. Each grey dot represents a pair of values of incremental cost and incremental effectiveness and the red point is the mean ICER points for each vaccine comparison. The grey shaded area below the diagonal cost-effectiveness threshold line ( $k=919.11$  USD) shows the cost-effective region



**Figure 3:** Cost-effectiveness acceptability curves showing the probability that each index scenario is cost-effective compared to the comparator over a range of cost-effectiveness thresholds (k=cost-effectiveness thresholds, pr=probability of cost-effectiveness)

## DISCUSSION

We assessed the epidemiological impact and cost-effectiveness of a range of COVID-19 vaccine deployment strategies and scenarios in Kenya. Our findings show that if Kenya had started with a full-scale vaccination programme in September 2021 and with an omicron-like variant introduced in November 2021, the deployment of COVID-19 vaccines in the Kenya population would likely avert a substantial number of cases, hospitalizations, and deaths from COVID-19. We find that a strategy to vaccinate mostly older adults (80% of those over 50y) who are at high risk of severe disease but which achieves low (30%) overall population coverage, yields the greatest reductions in severe infections and deaths per fully vaccinated adult. The marginal health benefits decrease with higher vaccine coverage levels (50% and 70%) as an increasing proportion of low risk younger adults, most with some immunity from previous infection, are vaccinated. These diminishing returns of increased coverage result in only the programme for older adults (i.e. the 30% coverage scenario) being cost effective while the expansion to younger age groups (i.e., 50% and 70%) was found not cost effective. Further, where an upsurge of SARS-CoV-2 occurs shortly after scale-up of vaccination (as modelled in this study) then deployment strategies that achieve rapid coverage of the target groups are more effective compared to slow vaccine deployment strategies.

Our findings are similar to evidence from South Africa, Madagascar, Pakistan, United Kingdom (UK), and United States of America (USA) that found vaccinating their population would decrease COVID-19 infections and deaths compared to a no vaccination scenario[8,37–40] and increasing vaccination coverage would increase the clinical benefits.[37,39] The South African study also found that a rapid vaccination roll-out pace resulted in “better” clinical outcomes (infections and deaths averted) and economic effectiveness compared to a non-rapid roll out pace.[37] The studies done in Madagascar, UK, and USA reported a greater impact when distribution of vaccines was prioritised according to the number of people of an older age in the region or among the elderly, reflecting similar findings to our study.[8,38,40] However, in contrast to the South African and Pakistan studies[37,39] who found that higher coverage scenarios had higher marginal impacts, we found that a minimal vaccine coverage of 30% of the adult Kenyan population targeting older age groups had the highest marginal impact. These differences could be explained by differences in the demographic profiles of the different populations of study. Higher population coverage with the COVID-19 vaccines have greater health impacts in countries that have higher proportions of the elderly and/or low previous exposure to COVID-19.

Using a societal perspective (that incorporates health system costs and productivity losses), we find that COVID-19 vaccination in Kenya is most cost-effective when targeted at older age groups in the population. This is because all our scenarios have the elderly covered first, and the incremental impact of increasing vaccination coverage among younger populations was less value for money. Given that the proportion of the elderly population in Kenya is low (11% of total population are aged 50 years and above),[13] targeting the COVID-19 vaccine to this vulnerable population achieves high cost-effectiveness at relatively low population-level vaccine coverage; 30% coverage of the population ensures that the maximal 80% of the older age group is vaccinated and a very low coverage of the younger age group (19%). Accounting for productivity losses improves the cost-effectiveness profile of COVID-19 vaccines, compared to when only direct health system costs are considered. For instance, for the 30% coverage scenarios with both a non-rapid and rapid deployment pace, the ICERs decreased on average by 342% and 652%, when the societal perspective was considered as opposed to the health system perspective, and as a result improving the cost-effectiveness profile. This underlines the limitations of using a narrow health system perspective that ignores broader societal costs of health system interventions. This is even more so for a vaccine deployed in a pandemic that has substantial socio-economic impacts, in addition to health impacts. These findings mirror cost-effectiveness studies



of COVID-19 vaccination done in Turkey and Pakistan that found that although COVID-19 vaccination strategies were cost-effective from a health system's perspective, they were cost-saving from a societal perspective.[39,41] This is in line with arguments from studies that estimate the public health value and impact of vaccination, which argue the need to broaden the perspectives for cost-effectiveness analysis of vaccines, as their impact is far-reaching, especially in the context of a pandemic.[42–44]

These findings have implications for COVID-19 vaccination policy in Kenya and other low-and middle-income countries (LMIC) settings with comparable demographic and COVID-19 epidemiological profiles. First, not unexpectedly, where an outbreak is imminent efforts to rapidly deploy the vaccine not only avert more cases, hospitalization, and deaths, but are also more cost-effective. By extension, had Kenya been able to deploy vaccines more rapidly, benefits would have been greater. Second, COVID-19 vaccination is likely to offer the best value for money when targeted to older age groups and possibly other vulnerable groups (such as those with risk increasing comorbidities) with high risk of severe disease and death, rather than to the whole population, in settings with overall low risk of severe disease and deaths, and high natural immunity due to previous exposure. This has several further implications. Kenya and other similar settings will achieve better health impacts and value for money with relatively small numbers of vaccines targeting the high-risk sections of the population. Against a context of constrained fiscal space for health, it is likely more prudent for Kenya and other African countries to target the vulnerable rather than whole populations. This consideration is likely to be even more relevant as African countries consider two shifts; the eligibility of children (below the age of 18 years) to COVID-19 vaccination and the transition to endemicity. If an endemic scenario will require annual vaccinations, Kenya and other African countries are unlikely to afford yearly vaccinations of their entire population. It is also apparent that such a strategy (vaccinating the entire population) is unlikely to be cost-effective, necessitating the need for Kenya and other African countries with comparable demographic and epidemic profiles to be both pragmatic and evidence-based in setting COVID-19 vaccine coverage policies and targets that are both feasible, effective, and cost-effective in their contexts (rather than replicating high income country strategies).

These results should be interpreted within the context of several limitations. First, our results are dependent on model assumptions and input parameters, as is the case with all modeling studies. We selected transmission model parameters based on published literature and available observation data. However, some data was limited, lacking, or uncertain and therefore we assumed our “best” estimate for Kenya. For example, we used estimates of vaccine effectiveness based on UK data and assume a duration of 14 days between vaccination and peak efficacy within our model structure. We noted from literature,[31–34] vaccine effectiveness varies with age, duration between vaccination and testing of efficacy, variant of infection, and the type of vaccine amongst other factors. The model does not consider the different professions of the population such as essential workers (health care workers, teachers, among others) as it focusses on age as the key risk group. However, frontline workers may be important to target since preventing infection among them lessens the potential impact on health and learning capacity. The latter might become more influential in the future with new vaccines if they are more effective in preventing re-infection and mild symptoms than current generation of vaccines. Second, sub-Saharan African countries like Kenya have notably reported lower cases and deaths compared to other countries across the globe, this could be attributed to their lower testing capacity. Hence, we assumed an under-reporting factor of 1:5 in hospitalised cases and deaths. Third, we instituted vaccination roll out near in time to the introduction of a new variant which enhances the benefit of rapid over slow roll-out. Distance between vaccine introduction and the emergence of an immune escape variant is likely to favour slower vaccine roll out. Fourth, assumptions about waning

immunity (natural and vaccine) and varying protection depending on variants affect the results. Fifth, in relation to the economic evaluation, although the cost-effectiveness analysis was conducted from a societal perspective, some costs have not been fully captured due to unavailability of data. These costs include household indirect costs incurred due to COVID-19 illness (e.g transport costs), costs as a result of long-COVID, and reduced productivity for those in the formal sector with asymptomatic/mild disease. In the latter, although we assume that they can resume work from home/places of quarantine they may have reduced productivity which is not captured in this analysis. These costs not captured in the analysis are however expected to be minimal. Fifth, the analysis assumed similar vaccine delivery costs for both rapid and non-rapid vaccination across similar coverage levels. However, it is likely that the rapid vaccination scenario may need more resources, especially cold chain equipment to hold larger batches of vaccines at a time. Sixth, the reported uncertainty of the ICER likely does not capture the full extent of the uncertainty, given the uncertainty of the costs of a yet to be established adult vaccination programme in Kenya. Lastly, the economic evaluation considers a 1.5-year time frame, potentially excluding costs and benefits of COVID-19 that may accrue over a longer period of time.

## **CONCLUSION**

This study contributes to the growing body of literature on the health impact and cost-effectiveness of COVID-19 vaccines. Kenya will achieve both greater marginal health impacts and better value for money if it prioritizes a targeted vaccination strategy among those at increased risk of severe disease and at a rapid rollout speed. The cost-effectiveness of the COVID-19 vaccine should be considered alongside other priority setting considerations in the Kenyan context.



## STATEMENTS

**Author Contribution:** EB, DJN, and PB conceptualised the study. SO, JO, CO, AK, RA, MO, AA, GMW, SU, EO, LIO, CAN, KK, PA, MM, RA, WN, IMO, JAGS, PB, and EB acquired the data. SPCB, JO, SO and CO analysed the data. SO, JO, SPCB, CO, PB, DJN, and EB were involved in the interpretation of the data and work. SO and JO wrote the first draft of the manuscript which was subsequently revised for important intellectual content by all authors. All authors read and approved the final manuscript.

**Funding:** This work was supported by funding from the Bill and Melinda Gates Foundation funded International Decision Support Initiative (IDSI), and funding from the National Institute for Health Research (NIHR) Global Health Research Unit on the Application of Genomics and Modelling to the Control of Virus Pathogens (17/63/82), on Mucosal Pathogens (16/136/46), and on Tackling Infections to Benefit Africa (16/136/33), using UK aid from the UK Government to support global health research, The UK Foreign, Commonwealth and Development Office and Wellcome Trust (grant# 220985/Z/20/Z).

**Acknowledgement:** We would like to acknowledge the Health Intervention and Technology Assessment Program (HITAP) and the Africa CDC Health Economic Program on their advice on the analysis.

**Competing interests:** None declared.

**Ethics approval:** The KEMRI Scientific and Ethics Review Unit approved this study under KEMRI/SERU/CGMR-C/4244.

**Patient consent:** Not required.

**Data sharing statement:** All code and data for the transmission model and economic evaluation analysis underlying this study is accessible at the Github repository: <https://github.com/SamuelBrand1/KenyaCoVaccines>.

## SUPPLEMENTARY INFORMATION

### Transmission model

#### Transmission model overview

The dynamics of SARS-CoV-2 transmission in each of the 47 Kenyan counties were assumed to follow a dynamic model adapted from a previous model;<sup>[4]</sup> here we extend the (modified SEIRS type) transmission model structure to include age stratification and vaccination. The epidemiological dynamics in each county are described by the following system of differential equations:

$$\begin{aligned}
 \frac{dS_{a,v}}{dt} &= -\sigma_a S_{a,v} (1 - \pi_v^{sus}) \lambda_a + \rho_{a,v-1} S_{a,v-1} - \rho_{a,v} S_{a,v} \\
 \frac{dE_{a,v}}{dt} &= \sigma_a (S_{a,v} + \sigma_{\omega 1} W1_{a,v} + \sigma_{\omega 2} W2_{a,v}) (1 - \pi_v^{sus}) \lambda_a - \alpha E_{a,v} + \rho_{a,v-1} E_{a,v-1} - \rho_{a,v} E_{a,v} \\
 \frac{dA_{a,v}}{dt} &= \alpha E_{a,v} (1 - \delta_a) - \gamma_A A_{a,v} + \rho_{a,v-1} A_{a,v-1} - \rho_{a,v} A_{a,v} \\
 \frac{dP_{a,v}}{dt} &= \alpha E_{a,v} \delta_a - \alpha_P P_{a,v} + \rho_{a,v-1} P_{a,v-1} - \rho_{a,v} P_{a,v} \\
 \frac{dD_{a,v}}{dt} &= \alpha_P P_{a,v} - \gamma_D D_{a,v} + \rho_{a,v-1} D_{a,v-1} - \rho_{a,v} D_{a,v} \\
 \frac{dR_{a,v}}{dt} &= \gamma_A A_{a,v} + \gamma_M M_{a,v} + \gamma_V V_{a,v} - \omega R_{a,v} + \rho_{a,v-1} R_{a,v-1} - \rho_{a,v} R_{a,v} \\
 \frac{dW1_{a,v}}{dt} &= \omega_1 R_{a,v} - \sigma_{\omega} \sigma_a W1_{a,v} (1 - \pi_v^{sus}) \lambda_a - \omega_2 W1_{a,v} + \rho_{a,v-1} W1_{a,v-1} - \rho_{a,v} W1_{a,v} \\
 \frac{dW2_{a,v}}{dt} &= \omega_2 W1_{a,v} - \sigma_a W2_{a,v} (1 - \pi_v^{sus}) \lambda_a + \rho_{a,v-1} W2_{a,v-1} - \rho_{a,v} W2_{a,v}
 \end{aligned} \tag{1}$$

The state variables are: Susceptible (**S**), latently infected (**E**), asymptotically infectious (**A**), pre-symptomatic infectious (**P**), symptomatic/diseased infectious (**D**), recovered and temporarily immune (**R**), and previously recovered/immune whose immunity to reinfection has waned (**W1**) before disappearing (**W2**). The indexing variables are age ( $a$ ) and vaccination dose status ( $v$ ), see below for further details on index variable structure. The age-specific force of infection was denoted  $\lambda_a$  (see below for details). The rate of progressing through latent uninfected stage ( $\alpha$ ), pre-symptomatic infectious state ( $\alpha_P$ ), rate of loss of complete immunity in two stages ( $\omega_1, \omega_2$ ), and the decreased susceptibility due to prior infection in stage **W1** ( $\sigma_{\omega}$ ), were assumed to be identical for all age groups, vaccination statuses, prior infection events, and eventual severity of the infection episode. The recovery rates ( $\gamma_A, \gamma_D$ ) depended on the severity of the episode, but not other factors. The baseline susceptibility per infectious contact ( $\sigma_a$ ) and probability of developing symptoms ( $\delta_a$ ), were assumed to depend on the age of the individuals.  $\pi_v^{sus}$  gives the effectiveness of vaccine status  $v$  at blocking transmission per infectious contact (see below). The per-capita rate at which a person in age group  $a$  and with vaccine status  $v$  transitioned to their next vaccine status was denoted  $\rho_{a,v}$  (see below). Figure S1 gives a visual overview of the model dynamics.

#### Age structure for transmission model

In this paper we used a coarse-grained set of age indices, partly to reduce number of model compartments and thereby increase the efficiency of parameter estimation (see below), and also

partly because the data used in parameter estimation did not support a finer grained age index. The six age groups corresponding to the age indices  $a = 1, \dots, 6$  were 0-19, 20-49, 50-59, 60-69, 70-79 and 80+ year olds. When using more fine-grained age-specific data we used population weighted average values over the finer age group categories within each of the coarser six age groups used in this model to generate data. The data we used in the transmission model which was translated from a larger number of age groups to the six used in this model was as following:

- Age- and setting- specific contact rates for Kenya from Prem and Jit.[45] These are provided in 16 5-year age groups and 75+ year olds. (Fig. S2)
- Age-specific relative susceptibility to infection ( $\sigma_a$ ). These were accessed from UK focused modelling and analysis.[8] (Fig. S3)
- Age-specific chance of a symptomatic episode given infection ( $\delta_a$ ). These were accessed from UK focused modelling and analysis.[8] (Fig. S3)

### Vaccine status and vaccination effects on transmission model dynamics and waning vaccination dynamics

We used five vaccination statuses to index individuals: unvaccinated ( $v = 1$ ), vaccinated with one ( $v = 2$ ) or two doses ( $v = 3$ ) with sufficient time elapsed since dose inoculation that maximum vaccine efficacy had been achieved (assumed to be 14 days), and two stages of waned vaccination ( $v = 4, 5$ ). The vaccination waning dynamics follow those described by Keeling et al.[16]

Vaccine status acted on the dynamics of the model in four distinct ways:

- 1) Decreasing the chance of infection per infectious contact by a factor  $1 - \pi_v^{sus}$  where  $\pi_v^{sus}$  was the effectiveness of vaccine dose status  $v$  at reducing infection.
- 2) Decreasing the probability of severe disease after infection by a factor  $1 - \pi_v^{dis}$  where  $\pi_v^{dis}$  was the effectiveness of vaccine dose status  $v$  against severe disease.
- 3) Decreasing the probability of death after infection by a factor  $1 - \pi_v^{death}$  where  $\pi_v^{death}$  was the effectiveness of vaccine dose status  $v$  against death.
- 4) Decreasing the infectiousness of dosed infecteds by a factor  $1 - \pi_v^{inf}$  where  $\pi_v^{inf}$  was the effectiveness of vaccine dose status  $v$  against transmitting infection.

We follow the “VE  $\rightarrow$  0%” scenario from Keeling et al,[16] where the effectiveness of the vaccine against acquisition of COVID-19 and infectiousness during a COVID-19 episode eventually decreases to zero ( $\pi_5^{sus} = 0, \pi_5^{inf} = 0$ ), whilst the eventual effectiveness against severe disease and death decreases to 70% ( $\pi_5^{dis} = \pi_5^{death} = 0.7$ ). The pre-waned vaccination status ( $v = 4$ ) has the same vaccine effectiveness as full second dose vaccination status ( $v = 3$ ) but is used to give non-exponential waning rates over an average of 430 days for both stages of waning vaccine effectiveness. Therefore, the per-capita transition rates  $\rho_{a,v}$  in equation (1) divide into:

$$\rho_{a,v} = \begin{cases} 0, & v = 0 \\ v_{a,v}, & v = 1, 2 \\ \xi_1, & v = 3 \\ \xi_2, & v = 4 \\ 0, & v = 5 \end{cases} . \quad (2)$$

Where  $v_{a,1}$  and  $v_{a,2}$  are the per-capita daily rates at which individuals in age group  $a$  receive their first and second doses of vaccine, and  $\xi_1 = \xi_2 = 2/430$  per day are the waning immunity rates of vaccine effectiveness.

Although, this model choice closely follows Keeling et al,[16] it should be noted that unlike Keeling et al we assume that protection from vaccines *and* natural infection combine favorably whereas in Keeling et al it was assumed that protection from natural infection dominated vaccine protection.

### Force of infection for transmission model

We define the age-dependent force of infection ( $\lambda_a$ ) in three steps. First, we define the effective number of infected in each age group  $b$ ,

$$I_b = \sum_v (1 - \pi_v^{inf})(\epsilon A_{b,v} + P_{b,v} + D_{b,v}) + I_{ext,b}. \quad (3)$$

The effective number of infecteds is the total number rescaled by decreased levels of infectiousness, such as the relatively lower infectiousness of asymptomatic infecteds compared to pre- and post-symptomatic infecteds ( $\epsilon$ ).  $I_{ext,b}$  represented an external coupling with infectious people external to internal transmission dynamics. We chose  $I_{ext,b}$  such that  $\sum_b I_{ext,b} = 100$ . Second, the rate of infectious contacts from an effective infected in age group  $b$  to anyone in age group  $a$  was defined as,

$$T_{ab} = \frac{\beta_0(t)(\beta_{home} T_{ab}^{home} + \beta_{school}(t) T_{ab}^{school} + \beta_{work} T_{ab}^{work} + \beta_{other} T_{ab}^{other})}{N_a}. \quad (4)$$

Contacts occur in any of four main settings: at home, at school, at work or in some other social setting, each of which has an age-specific contact rate matrix for Kenya ( $T_{ab}^{home}, T_{ab}^{school}, T_{ab}^{work}, T_{ab}^{other}$ ) estimated by Prem and Jit.[45] Google mobility data[46] and previous epidemic modelling in Kenya [4] suggest that contact rates in Kenya had returned to approximately pre-pandemic baseline by January 2021, therefore, we treated the setting specific transmission rates per contact for at home, at work and other social setting ( $\beta_{home}, \beta_{work}, \beta_{other}$ ) as constant. Transmission rate per contact at schools ( $\beta_{school}(t)$ ) was constant during term time, but dropped to zero during Kenyan school holidays (19<sup>th</sup> March – 10<sup>th</sup> May, 16<sup>th</sup> July – 26<sup>th</sup> July, 1<sup>st</sup> October – 11<sup>th</sup> October, 23<sup>rd</sup> December – 4<sup>th</sup> January). The baseline transmission rate per contact ( $\beta_0(t)$ ) varied according to the SARS-CoV-2 variant frequency in the county (see below). Third, the force of infection was defined as,

$$\lambda_a = \sum_b T_{ab} I_b. \quad (5)$$

### Alpha, Beta and Delta variant frequency effect on transmission model dynamics

The Alpha, Beta and Delta variants of SARS-CoV-2 have circulated in Kenya,[47] displaying a pattern of sequential dominance. From December 2020, Alpha and Beta variants increased in frequency ([48] Fig. S2), with a complex spatial pattern of relative frequency of Alpha vs Beta within Kenya.[48] Then, from May 2021, the frequency of Delta variant increased very rapidly to complete domination across Kenya (Fig. S2).

We modelled the effect of variant frequency on the baseline transmission rate per infectious contact  $\beta_0(t)$  as a sequence of strain dominations each occurring over a timescale set by a logistic growth curve,

$$\beta_0(t) = \beta_{wt} [1 + R_{\alpha\beta} \text{logistic}(t; r_{\alpha\beta}, T_{\alpha\beta})] [1 + R_{\delta} \text{logistic}(t; r_{\delta}, T_{\delta})]. \quad (6)$$

Where  $\text{logistic}(t; r, T) = \frac{\exp(r(t-T))}{1 + \exp(r(t-T))}$ . In this model,  $\beta_{wt}$  is the baseline transmission rate per infectious contact of the origin strain(s), called wild-type strains, circulating in Kenya before Alpha and Beta variants,  $(1 + R_{\alpha\beta})$  is the proportional change in the reproductive number for SARS-CoV-2 after domination by Alpha or Beta variant relative to wild-type strains, and  $(1 + R_{\delta})$  is the proportional change in the reproductive number for SARS-CoV-2 after domination by Delta variant relative to Alpha or Beta variant.  $r_{\alpha\beta}, r_{\delta}, T_{\alpha\beta}, T_{\delta}$  set the exponential rates and timing of the logistic growth curves. The implied relative frequencies of wild-type, Alpha/Beta and Delta variants over time are,

$$\begin{aligned} f_{wt}(t) &= [1 - \text{logistic}(t; r_{\alpha\beta}, T_{\alpha\beta})] [1 - \text{logistic}(t; r_{\delta}, T_{\delta})], \\ f_{\alpha\beta}(t) &= \text{logistic}(t; r_{\alpha\beta}, T_{\alpha\beta}) [1 - \text{logistic}(t; r_{\delta}, T_{\delta})], \\ f_{\delta}(t) &= \text{logistic}(t; r_{\delta}, T_{\delta}). \end{aligned} \quad (7)$$

### Transmission model observables: Proportions PCR test and serology test positive

The underlying transmission of SARS-CoV-2 is not observed, rather we have access to swab tests and serological tests (positive and negative) aggregated by date, age, and county. The chance infected individuals test positive for either type of test depends on the number of days post-infection. Therefore, we coupled the dynamics cumulative infections to transmission model,

$$\begin{aligned} \frac{dF_{a,v}}{dt} &= \sigma_a S_{a,v} (1 - \pi_v^{sus}) \lambda_a, \\ \frac{dC_{a,v}}{dt} &= \sigma_a (S_{a,v} + \sigma_{\omega} W1_{a,v} + W2_{a,v}) (1 - \pi_v^{sus}) \lambda_a, \\ \frac{d\bar{C}_{a,v}}{dt} &= \sigma_a (S_{a,v} + 0.5\sigma_{\omega} (1 - \pi_5^{dis}) W1_{a,v} + (1 - \pi_5^{dis}) W2_{a,v}) (1 - \pi_v^{sus}) \lambda_a. \end{aligned} \quad (8)$$

Where  $F_{a,v}(t)$  and  $C_{a,v}(t)$  were the cumulative number of people in the county in age group  $a$  and vaccine status  $v$  infected by time  $t$ , respectively split by it being their first infection episode or any infection episode.  $\bar{C}_{a,v}(t)$  was the reinfection disease risk-weighted cumulative infection rate; that is the total infections discounted by the decreased risk of severe disease among reinfections (see below for use of this observable). In the absence of other evidence, we follow Keeling et al[16] in assuming that protection against disease due to prior infection is similar to that of vaccination. For the fully waned immunity post-natural infection state (**W2**) we assume that the protection from disease is equivalent to fully waned vaccine protection, and, for the partially waned immunity post-natural infection state (**W1**) that this protection is halved relative to **W2**.

The daily new infections, on each day  $n$  starting at  $t = n$  are then,

$$\begin{aligned} \iota_{F,a,v}(n) &= F_{a,v}(n+1) - F_{a,v}(n), \quad n = 1, 2, \dots \\ \iota_{C,a,v}(n) &= C_{a,v}(n+1) - C_{a,v}(n), \quad n = 1, 2, \dots \end{aligned} \quad (9)$$

The probability that an infected individual would be determined as having been infected  $\tau$  days after infection if tested by either a PCR swab test or a serology test were denoted, respectively,  $Q_{PCR}(\tau)$  and  $Q_{sero}(\tau)$ . We used the same  $Q_{PCR}$  and  $Q_{sero}$  probabilities as in Brand et al 2021.[4]

By combining the underlying infection processes and the delay between infection and observability in our available data sets we find that the number of people who would test positive on each day  $n$  in each county with either a PCR test ( $P_a^+(n)$ ), or a serology test ( $S_a^+(n)$ ), was,

$$\begin{aligned} P_a^+(n) &= \sum_{s=1}^{n-1} \sum_v \iota_{C,a,v}(s) Q_{PCR}(n-s), \\ S_a^+(n) &= \sum_{s=1}^{n-1} \iota_{F,a,1}(s) Q_{sero}(n-s) + p_{FP} \sum_v S_{a,v}(t). \end{aligned} \tag{10}$$

where  $t$  was the midpoint of day  $n$ .

$p_{FP}$  was the false positive rate for the serology assay (see table S1). Underlying assumptions for equation (10) are: 1) that the PCR test is 100% specific to SARS-CoV-2, 2) that only the first infection contributes to the serological status of individuals, but that reinfections contribute to PCR status equally to first infections, and 3) that during the period of transmission parameter inference there were effectively zero vaccinated individuals, and therefore, all seropositivity was evidence of prior nature infection. We do not use any serological data from Kenya after May 2021 when the vaccination rate was very low in Kenya, and, therefore, vaccinations have negligible effect on prevalence of SARS-CoV-2 specific antibodies.

The number of people PCR positive is not observed directly, but rather test positive and negative swab test samples. We consider the proportion of these daily samples that are positive to be a potentially biased sample of the true underlying proportion that would be PCR-positive if everyone was tested ( $\sum_a P_a^+ / \sum_a N_a$ ). Therefore, we model the expected proportion PCR test positive (over all age groups) on day  $n$  as,

$$\bar{P}^+(n) = \frac{\chi \sum_a P_a^+(n)}{(\chi - 1) \sum_a P_a^+(n) + N}. \tag{11}$$

Where  $\chi$  is an observed swab sample bias parameter, where  $\chi = 1$  indicates unbiased sampling,  $\chi < 1$  indicates bias in favour finding PCR negative individuals (i.e.,  $\bar{P}^+(n) < \sum_a P_a^+ / \sum_a N_a$ ), and  $\chi > 1$  indicates bias in favour of finding PCR positive individuals (i.e.,  $\bar{P}^+(n) > \sum_a P_a^+ / \sum_a N_a$ ).

## Clinical outcome model

### Clinical outcomes of SARS-CoV-2 infections

Severe infections eventually lead to a clinical outcome. We consider three possibilities in this model:

- 1) **Deadly outcome.** Deadly infected individuals die after a delay period defined by the probability distribution  $f_\mu$ . Death, conditional on infection, occurs with probability  $\mu_a^D$ .
- 2) **Critical outcome.** Critically infected individuals require a stay in ICU for a duration defined by the probability distribution  $f_{ICU}$ , then move to a general ward in a hospital or health facility, where they stay for a duration defined by the probability duration  $f_{hosp}$ . Critical disease, conditional on infection, occurs with probability  $\mu_a^C$ .



- 3) **Severe outcome.** Severely infected individuals require a stay in a general ward in a hospital or health facility, where they stay for a duration defined by the probability duration  $f_{hosp}$ . Severe disease, conditional on infection, occurs with probability  $\mu_a^S$ .

**Clinical outcome model observables: Reported incidence of deaths, occupancy of general wards and Intensive care units**

Incidence rate of clinical outcomes. The lag between infection and needing treatment, for those infected individuals who die, was defined as the convolution of two time-duration distributions:

1. The duration of time between infection and symptoms (days), which we assumed was distributed  $LogNormal(\widehat{log\mu} = 1.64, \widehat{logstd} = 0.36)$ . [49]
2. The duration of time between initial symptoms and severe symptoms (days), sufficient to seek hospitalisation, which we assumed was distributed  $U(1,5)$ . [50]

We discretized the two distributions to give probability functions  $f_{IS}$  for the number of days between infection and symptoms, and  $f_{SH}$  for the number of days between symptom onset and severe symptom onset. The probability function for the (discrete) number of days between infection and severe or critical disease, for those who died,  $f_{dis}$ , was given as a discrete convolution over these probability mass functions:

$$f_{dis}(\tau) = [f_{IS} * f_{SH}](\tau) \text{ for the probability that severe or critical disease leads to seeking medical assistance } \tau \text{ days after infection, conditional on that as outcome.} \quad (12)$$

We use this delay distribution to give a rate of people in age group  $a$  requiring medical treatment on day  $n$  up to some unknown age-dependent and variant-dependent risk-factor, which will fit against available data on reported severe, critical and deadly outcomes,

$$t_{dis,a,v,r}(n) \propto \sum_{s < n} f_r(s) [\bar{C}_{a,v}(s+1) - \bar{C}_{a,v}(s)] f_{dis}(n-s). \quad (13)$$

Where  $f_r(s)$  is the relative frequency of variant  $r$  on day  $s$  (see equation (7)). Note that in equation (13) the reduction in risk due to reinfection is already accounted for (see equation (8)).

Observation of incidence of deadly outcome of infection. There is likely to be under-reporting of deaths due to COVID-19 in Kenya. [35,51] In this paper, we don't have sufficient data to estimate the true level of under-reporting of deaths in Kenya. However, by assuming that the age-dependent risk of death after infection with SARS-CoV-2 is the same in every county, we can estimate county-specific under-reporting/change in risk *relative* to the capital Nairobi. Concretely, we model the expected number of observed deaths on each day  $n$ , in each age group  $a$ , and each county  $c$ , as,

$$Deaths_{a,c}(n) = \psi_c \mu_a^D \sum_{s < n} \sum_v \sum_r \psi_r^D (1 - \pi_v^{death}) t_{dis,a,v,r}(s) f_\mu(n-s). \quad (14)$$

Where  $\psi_c$  is the county-specific under-reporting rate with  $\psi_{nairobi} = 1$ ,  $\psi_r^D$  is the relative risk of death by variant with  $\psi_{wt}^D = 1$ , and the duration of time between needing treatment and death had probability distribution  $f_\mu$ .

Observation of hospital and ICU occupancy due to severe or critical infections. We don't have access to reports on the incidence of severe and critical cases arriving at hospitals/health facilities and ICUs; the relevant observables from the clinical outcome model are occupancies of patients overall in Kenya by setting rather than arrival of patients at those settings.

The probability distributions for length of stay in either hospital/health facility or ICU are denoted as  $f_{hosp}$  and  $f_{ICU}$ , implying upper distribution functions  $Q_{hosp}(n) = \sum_{s>n} f_{hosp}(s)$  and  $Q_{ICU}(n) = \sum_{s>n} f_{ICU}(s)$  for the probability that a stay in hospital or ICU is longer than  $n$  days. The expected total number of people in intensive care units on day  $n$  is,

$$ICU(n) = \sum_{a,c} \psi_c \mu_a^c \sum_{s<n} \sum_v \sum_r \psi_r^c (1 - \pi_v^{dis}) \iota_{dis,a,v,r}(s) Q_{ICU}(n-s). \quad (15)$$

Where  $\psi_r^c$  is the relative risk of critical disease by variant with  $\psi_{wt}^c = 1$ . After a critical case has completed a stay in an ICU, we model them as having a stay in a general ward with the same distribution of length as per a severe case admitted to a general ward (without a stay in ICU). The upper distribution function for the whole stay in ICU and general ward is then  $Q_{ICUH}(n) = \sum_{s>n} [f_{ICU} * f_{hosp}](s)$ , giving the expected number of critical cases in either ICU or general ward as,

$$ICUH(n) = \sum_{a,c} \psi_c \mu_a^c \sum_{s<n} \sum_v \sum_r \psi_r^c (1 - \pi_v^{dis}) \iota_{dis,a,v,r}(s) Q_{ICUH}(n-s). \quad (16)$$

The expected number of patients occupying general wards is the addition of severe cases who have been admitted directly to general wards, and critical cases who have completed their stay in ICU and are now in general wards,

$$HOSP(n) = [ICUH(n) - ICU(n)] + \sum_{a,c} \psi_c \mu_a^s \sum_{s<n} \sum_v \psi_r^s (1 - \pi_v^{dis}) \iota_{dis,a,v,r}(s) Q_{hosp}(n-s). \quad (17)$$

Where  $\psi_r^s$  is the relative risk of severe disease by variant with  $\psi_{wt}^s = 1$ .

## Parameter Inference

In this work, we make inferences on two groups of parameters:

1. The parameters of the transmission model in [equations (1-5)], and the bias parameter ( $\chi$ ) for the observed versus actual proportion PCR positive in daily swab test [equation (11)].
2. The parameters of the clinical outcome model [equations (12-15)]: the relative under-reporting rate by Kenyan county  $\psi_c$ , the age-dependent clinical outcome probabilities  $\mu_a^s, \mu_a^c$ ,  $\mu_a^D$ , and the variant specific clinical outcome probabilities  $\psi_r^D, \psi_r^c, \psi_r^s$ .

For the transmission model parameters, we used Bayesian inference to infer a joint posterior distribution for the parameters for each county. For the clinical outcome model, we inferred parameters by minimizing the divergence between model prediction of the observables [equations (14-15) and (17)] and actual reporting, under the assumption that the rate of people arriving for medical treatment, up to the unknown risk factors [equation (13)], was that implied by the posterior mean prediction implied by the Bayesian inference of the transmission model parameters.

A challenge with using the linelist data in Kenya for inference of transmission was that the metadata concerning the reason for receiving a swab test, the levels of symptoms of people who tested positive, and their healthcare outcomes were often missing. Overall, more than 90% of the people who tested positive in Kenya, and for whom we have a description of their symptoms, reported no symptoms (asymptomatic). Therefore, unlike model-based inference for COVID-19 transmission in high-income

countries we didn't use severe outcomes such as hospitalization or death as data sources for inference, e.g., [9,52,53] because this data was unreliable. Instead, we concentrated on fitting to the proportion positive of daily swabs test and serological tests jointly with detection rate of cases (see **Transmission model observables: Proportions PCR test and serology test positive** above). It should be noted that this meant that we didn't use age-specific PCR test data, but rather fitted to the aggregate proportion positive over all age groups. However, we did use age-specific seroprevalence data.

We describe the three main ingredients for our Bayesian approach below: 1) the log-likelihood function for the data given a set of parameters, 2) the county-specific hierarchy of prior distributions for the parameters, and, 3) the Markov-chain Monte Carlo method used to draw parameter sets from the posterior distribution.

### Bayesian inference of transmission model parameters

Data and log-likelihood function for transmission model. Given the daily PCR and serology data for a county,

$$\mathcal{D}_c = \{ObsS_a^+(n), ObsS_a^-(n), ObsP^+(n), ObsP^-(n)\}_{a=1,\dots,6; n=1,2,3,\dots} \quad (18)$$

The log-likelihood function for the unknown transmission parameters ( $\theta_{TM}$ ) in that county was,

$$l(\theta_{TM}) = P(\mathcal{D}_c | \theta_{TM}) = \sum_n \ln f_{BB}(ObsP^+(n) | \widehat{N}_s = N_{PCR}(n), \hat{p} = \bar{P}^+(n), \widehat{M} = M_{PCR}) \\ + \sum_{n,a} \ln f_{BB}(ObsS_a^+(n) | \widehat{N}_s = N_{sero,a}(n), \hat{p} = \frac{S_a^+(n)}{N_a}, \widehat{M} = M_{sero}). \quad (19)$$

Where  $f_{BB}(x | \widehat{N}_s, \hat{p}, \widehat{M})$  is the probability function for a Beta-binomial with sample size  $\widehat{N}_s$ , expected proportion of successes  $\hat{p}$ , and effective sample size  $\widehat{M}$ . This was a convenient reparameterization of the Beta-binomial model under the transformation  $\widehat{N}_s = n, \hat{p} = \frac{\alpha}{\alpha + \beta}, \widehat{M} = \alpha + \beta$  from the typical  $(n, \alpha, \beta)$  parameterization.  $N_{PCR}(n)$  and  $N_{sero,a}(n)$  were the total number of samples (positive and negative) of, respectively, PCR test and serological tests on day  $n$ , in age group  $a$ .  $\bar{P}^+(n)$  was derived from the transmission model for each day  $n$  as per above [equations (10-11)].  $M_{PCR}$  and  $M_{sero}$  were fixed from a previous modelling study.[4] The first day where samples were included in the log-likelihood calculation was 1st January 2021.

Initial conditions and model simulation. We fit to data from 1<sup>st</sup> January 2021, however, because PCR cases and serological detection are lagged indicators of infections weeks previously, we start the model simulation in each county on 1<sup>st</sup> December 2020 and use the first month of simulation to allow the simulation to converge onto the epidemic dynamics. Simulation of the model was done by solving the ODE system [equation (1)] forwards from the county-specific initial conditions, with county-specific parameter configuration, using an explicit/implicit switching solver provided by the **DifferentialEquations.jl** Julia programming language package.[54]

We reduced the number of unknown parameters for the initial state of the epidemic model by considering only the overall latent infected numbers ( $E_0$ ) and a scale factor on the proportion exposed to COVID (i.e. in  $R/W1/W2$  epidemiological compartments) relative to a cross-sectional survey done in Nairobi in mid-November 2020[55] which we denote  $\tau$ . We fix the initial removed and waned immunity numbers as

$$\begin{aligned} R_{a,1}(0) &= N_a(0) * P_{sero,a} * \tau * 0.95, \\ W1_{a,1}(0) &= N_a(0) * P_{sero,a} * \tau * 0.05, \\ W2_{a,1}(0) &= 0. \end{aligned} \quad (20)$$

Where  $P_{sero,a}$  was the raw (non-test sensitivity adjusted) seroprevalence estimate for Nairobi in the cross-sectional survey,[55] and  $\tau$  was an adjustment factor which was added to the set of parameters to be inferred in the set  $\theta_{TM}$ . Note that we are assuming that 5% of the previously exposed population have lost complete immunity to reinfection by 1<sup>st</sup> December 2020, and that no previously exposed people had completely lost immunity to reinfection. The adjustment factor  $\tau$  allowed the model flexibility to represent counties with lower seroprevalence data in 2021 as having had a smaller initially exposed fraction compared to Nairobi, whilst also allowing upwards adjustment to account for the fact that the second wave of cases in Kenya occurred during the Nairobi cross-sectional study.

The age-specific numbers of initially latent infected people were derived from the next-generation matrix  $K(P_{sero,1}, \dots, P_{sero,6}, \tau, \beta_0, \beta_{school}, \beta_{home}, \beta_{other}, \beta_{work})$ , where we have made explicit the parameters being inferred that the next-generation matrix depends upon and its explicit dependence on the baseline seroprevalence estimates. The eigenvector  $\mathbf{v}$ , normalized such that  $|\mathbf{v}|_1 = 1$ , associated with the leading eigenvalue of  $K$ , represented the expected distribution of new infections across age groups. Therefore, we specified

$$\mathbf{E}(0) = E_0 \mathbf{v}. \quad (21)$$

Where  $\mathbf{E}(0) = [E_{1,0}(0), \dots, E_{6,0}(0)]^T$ . The rest of the initial variables were specified as being dependent of the flow out of the latent infected state,

$$\begin{aligned} A_{a,0}(0) &= \alpha * E_{a,0}(0) * (1 - \delta_a)/(1 + \gamma_A), \\ P_{a,0}(0) &= \alpha * A_{a,0}(0) * \delta_a/(1 + \alpha_p), \\ D_{a,0}(0) &= \alpha_p * P_{a,0}(0) * (1 - h_a)/(1 + \gamma_D). \end{aligned} \quad (22)$$

In every county, every individual in the model was initially unvaccinated.

Priors. In every county we used the following priors for parameter inference:

- $\epsilon \sim \text{Beta}(\hat{\alpha} = 50, \hat{\beta} = 50)$ .
- $\beta_0 \sim \Gamma(\hat{k} = 10, \hat{\theta} = 1.5/10)$ .
- $\beta_{other} \sim \Gamma(\hat{k} = 10, \hat{\theta} = 1.5/10)$ .
- $\beta_{home} \sim \Gamma(\hat{k} = 10, \hat{\theta} = 1.5/10)$ .
- $\beta_{work} \sim \Gamma(\hat{k} = 10, \hat{\theta} = 1.5/10)$ .
- $\beta_{school} \sim \Gamma(\hat{k} = 10, \hat{\theta} = 1.5/10)$ .
- $R_{\alpha\beta} \sim \Gamma(\hat{k} = 15, \hat{\theta} = \frac{0.4}{15})$ .
- $r_{\alpha\beta} \sim \Gamma(\hat{k} = 15, \hat{\theta} = \frac{0.15}{15})$ .
- $R_{\delta} \sim \Gamma(\hat{k} = 15, \hat{\theta} = \frac{0.6}{15})$ .
- $r_{\delta} \sim \Gamma(\hat{k} = 15, \hat{\theta} = \frac{0.2}{15})$ .

- $E_0 \sim \Gamma(\hat{k} = 3, \hat{\theta} = \frac{1000}{3})$ .
- $\tau \sim \Gamma(\hat{k} = 5, \hat{\theta} = \frac{1.5}{5})$ .

The prior for the PCR observation bias parameter  $\chi$ , differed between counties (see below). For Nairobi and Mombasa we used a prior:

- $\chi \sim \Gamma(\hat{k} = 3, \hat{\theta} = \frac{4.5}{3})$ .

MCMC draws. We used Hamiltonian MCMC with NUTS[56,57] to perform Bayesian inference by drawing 2,000 samples from the posterior distribution,

$$\theta_{TM}^{(k,c)} \sim P(\theta_{TM} | \mathcal{D}_c) \propto \exp(l(\theta_{TM})) \pi(\theta_{TM}), \text{ for } k = 1, 2, 3, \dots \quad (23)$$

for each county using the NUTS-HMC sampler implemented by the Julia language package *dynamicHMC.jl*. The HMC method required a log-likelihood gradient,  $\nabla_{\theta}(l + \pi)$ , which, for our use-case of an ODE system with a comparative low number of parameters (<100 parameters), was most efficiently supplied by forward-mode automatic differentiation implemented by the package *ForwardDiff.jl*. The MCMC chain converged for each county (all MCMC chains and MCMC diagnostics can be accessed through the linked open code repository: <https://github.com/SamuelBrand1/KenyaCoVaccines>). The posterior mean (and 95% CIs) for each parameter can also be found in the open code repository.

Approximate county-specific hierarchical model for PCR observation bias. The serological data is important for our inference because it gives information about the proportion of each age group infected at different time points, and, therefore, allows the PCR observation bias parameter ( $\chi$ ) to be identifiable. However, the amount of serological data differs from county to county. To allow cross-inference between counties for the bias parameter we assumed that 1) Nairobi and Mombasa were sufficiently distinct from other counties that the inferred bias parameter for these city/counties was not relevant to other counties, and 2) the other 45 counties had a bias parameter drawn from a common distribution,  $\chi_c \sim \Gamma(k_{\chi}, \theta_{\chi})$ , where  $k_{\chi}$  and  $\theta_{\chi}$  are the hyperparameters of this hierarchical model. This reflected our underlying belief that despite regional variations in transmission, the observation of data would be similar in all counties outside of the main two urban hubs.

A fully Bayesian approach to inference would involve including  $\{\chi_c\}_c$  and the hyperparameters  $k_{\chi}, \theta_{\chi}$  within a joint log-likelihood over all Kenyan counties (except Nairobi and Mombasa). However, to accelerate inference we used an approximation to this hierarchical model. The 9 counties with the most amount of serological data available, apart from Nairobi and Mombasa, were **Embu, Kilifi, Kisii, Kisumu, Kwale, Nakuru, Nyeri, Siaya, and Uasin Gishu**. We performed MCMC draws for each of these counties using a prior  $\chi \sim \Gamma(\hat{k} = 3, \hat{\theta} = \frac{4.5}{3})$ , which gathered a set of MCMC draws for  $\chi$  from the posterior distribution for each county  $c$ ,  $\{\chi^{(k,c)}\}_{k=1,\dots,2000} \sim P(\chi | \mathcal{D}_c)$ . We then approximated maximum a-posteriori (MAP) estimates for the hyperparameters  $k_{\chi}, \theta_{\chi}$  using,

$$\ln P(k_{\chi}, \theta_{\chi} | \mathcal{D}_1, \dots, \mathcal{D}_9) = \ln \pi(k_{\chi}, \theta_{\chi}) + \sum_c \ln \sum_k f_{\Gamma}(\chi^{(k,c)} | \hat{k} = k_{\chi}, \hat{\theta} = \theta_{\chi}) + \text{const.} \quad (24)$$

Where  $f_{\Gamma}$  is the density function of a Gamma distribution, and  $\ln \pi(k_{\chi}, \theta_{\chi})$  was the log-prior for the hyper-parameters. The hyper-priors used were:

- $k_{\chi} \sim \exp(\hat{\mu} = 10)$ .



- $\theta_\chi \sim \exp(\hat{\mu} = 0.45)$ .

The MAP estimation is done under the distributional assumption that the observed outcome data was distributed negative binomially with the same mean as the posterior predictive mean value (that is averaged over the posterior predictive distribution for the infection process), and a negative binomial clustering factor inferred jointly with the risk factors.

We then could use the 9 non-city counties with the most serological data to create MAP estimates  $\hat{k}_\chi$ ,  $\hat{\theta}_\chi$  by maximizing equation (24). Equation (24) represented an approximation where we treated the MCMC draws of the  $\chi$  parameter as “data” for making inference on the hyper-parameters despite using a different prior to generate the MCMC samples.

The second approximation is that for the 36 other counties that were not Nairobi, Mombasa, or one of the 9 listed above, we used a prior for  $\chi$  generated from these MAP estimates

- $\chi \sim \Gamma(\hat{k} = k_\chi/2, \hat{\theta} = 2\theta_\chi)$ .

The reason for the scaling of 2 was to increase the prior variance for  $\chi$  to reflect that we are approximating a hierarchical model.

### Minimum divergence estimates for the infection outcome model

After performing MCMC we were able to estimate the posterior mean for the rate of diseased incidence, up to a proportionality with unknown risk factors  $E[l_{dis,a,v,r}(n)|\mathcal{D}]$ , for each day  $n$ , age group  $a$ , variant  $v$ , and county  $c$ , by solving the ODE system [equation (1)] for each set of transmission model parameters drawn from the MCMC and using equation (13). We can then define the posterior expected number of deaths  $\overline{Deaths}_{a,c}(n; \psi_c, \psi_r^D, \mu_a^D)$  by replacing the parameter-specific diseased incidence rate in equation (14) with its posterior mean. We define the divergence due to a choice of age dependent mortality rates ( $\mu_a^D$ ), relative reporting rate ( $\psi_c$ ), relative variant-specific risk of death ( $\psi_r^D$ ) and clustering factor  $\alpha_D$ ,

$$Div_D(\psi_c, \psi_r^D, \mu_a^D, \alpha_D) = -2 \sum_{a,c,n} \ln f_{NB}(ObsDeaths_{a,c}(n) | \hat{\mu} = \overline{Deaths}_{a,c}(n; \psi_c, \psi_r^D, \mu_a^D), \hat{\alpha} = \alpha_D). \quad (25)$$

Where  $f_{NB}(x|\hat{\mu}, \hat{\alpha})$  is the probability function for a negative binomial with mean  $\hat{\mu}$  and clustering factor  $\hat{\alpha}$ , and  $ObsDeaths_{a,c}(n)$  are the daily reported deaths in each age group, each county and on each day. We found a minimum point for equation (25), which we used as estimators,  $\hat{\psi}_c, \hat{\psi}_v^D, \hat{\mu}_a^D, \hat{\alpha}_D$  (Figs. S2, S3).

The ICU occupancy data was not available broken down by age or county, therefore, we assumed that the risk of reported critical disease was proportional to the risk of death for everyone and focused on fitting the relative risk of critical disease vs death  $\mu_{rel}^{CD}$ . We defined the divergence between model prediction of ICU occupancy and observed occupancies due to relative risk of critical disease vs death  $\mu_{rel}^{CD}$ , variant specific risk of critical disease  $\psi_r^C$ , and clustering factor  $\alpha_C$ ,

$$Div_{ICU}(\mu_{rel}^{CD}, \psi_v^C, \alpha_C) = -2 \sum_n \ln f_{NB}(ObsICU(n) | \hat{\mu} = \sum_{a,c} \overline{ICU}_{a,c}(n; \hat{\psi}_c, \psi_r^C, \mu_{rel}^{CD}, \hat{\mu}_a^D), \hat{\alpha} = \alpha_C). \quad (26)$$

Where  $ObsICU(n)$  was the reported Kenyan National ICU occupancy on day  $n$ , and  $\overline{ICU}_{a,c}$  is the posterior mean value for the ICU occupancy with COVID. The minimum point for equation (26) gave estimators,  $\hat{\psi}_v^C, \hat{\mu}_a^C = \hat{\mu}_{rel}^{CD} \hat{\mu}_a^D, \hat{\alpha}_C$  (Fig. S2).

The general ward occupancy data was also not available broken down by age or county, therefore, we assumed that the risk of reported severe disease was proportional to the risk of death for everyone and focused on fitting the relative risk of severe disease vs death  $\mu_{rel}^{SD}$ . We defined the divergence between model prediction of general ward occupancy and observed occupancies due to relative risk of severe disease vs death  $\mu_{rel}^{SD}$ , variant specific risk of severe disease  $\psi_r^S$ , and clustering factor  $\alpha_S$ ,

$$Div_{HOSP}(\mu_{rel}^{SD}, \psi_r^S, \alpha_S) = -2 \sum_n \ln f_{NB}(ObsHOSP(n)|\hat{\mu} = \sum_{a,c} \overline{HOSP}_{a,c}(n; \hat{\psi}_c, \psi_r^S, \mu_{rel}^{SD}, \hat{\mu}_a^D), \hat{\alpha} = \alpha_S). \quad (27)$$

Where  $ObsHOSP(n)$  was the reported Kenyan National general ward occupancy with COVID on day  $n$ , and  $\overline{HOSP}_{a,c}$  is the posterior mean value for the ICU occupancy, with the contribution from patients arriving into general wards from ICU already calculated using the minimum divergence estimates from equation (26). The minimum point for equation (27) gave estimators,  $\hat{\psi}_r^S$ ,  $\hat{\mu}_a^S = \hat{\mu}_{rel}^{SD} \hat{\mu}_a^D$ ,  $\hat{\alpha}_S$  (Fig. S2).

### Vaccine scenario projections and immune-escape variant

Vaccination rates. We considered 7 vaccine rollout scenarios starting from 1<sup>st</sup> September 2021: No vaccination (baseline), 30%, 50%, 70% target coverage of Kenyan over 18s with either an 18 month or 6-month (rapid) time scale (Table 1). In each case we assumed that:

1. The number of vaccines deployed in each county each day was constant over the rollout.
2. Second dose followed first dose after a 56 day lag.
3. Over 50 year olds were offered the vaccine first, but demand saturated at 80% coverage among over 50 year olds (age groups 3-6 in the model) and afterwards the vaccine was offered to 18-49 year olds (age group 2).
4. Vaccines were deployed pro-rata across all disease/infection states; that is there was no dependence on past infection history in seeking vaccines.

Mathematically, this corresponds to this choice for the per-capita vaccination rate (equation (2)) for the first dose, that is pro-rata distribution among all unvaccinated groups in stages by age,

$$v_{a,1}(t) = \frac{NV_{scenario} \mathbf{1}(\sum_{a=3,\dots,6} \sum_{v=2,\dots,5} N_{a,v} < 0.8 \sum_{a=3,\dots,6} N_a)}{T_{scenario} \sum_{a=3,\dots,6} N_{a,1}}, \quad a = 3,4,5,6$$

$$v_{2,1}(t) = \frac{NV_{scenario} \mathbf{1}(\sum_{a=3,\dots,6} \sum_{v=2,\dots,5} N_{a,v} \geq 0.8 \sum_{a=3,\dots,6} N_a)}{T_{scenario} N_{2,1}}, \quad (28)$$

For  $1st\ sept\ 2021 \leq t \leq 1st\ sept\ 2021 + T_{scenario} - 56\ days$

Where  $N_{a,v}$  is the county population by age and vaccine status (summed over disease/infection status).  $\mathbf{1}(\cdot)$  is an indicator function enforcing that the vaccination rate among over 50s (age groups 3-6) drops to zero at an 80% coverage.  $NV_{scenario}$  is the number of doses implied by the scenario target coverage, and  $T_{scenario}$  is the time in days over which this target coverage is to be achieved in the scenario. The second dose per capita rate is like the first dose but with doses distributed among people who have had their first dose,

$$v_{a,2}(t) = \frac{NV_{scenario} \mathbf{1}(\sum_{a=3,\dots,6} \sum_{v=3,\dots,5} N_{a,v} < 0.8 \sum_{a=3,\dots,6} N_a)}{T_{scenario} \sum_{a=3,\dots,6} N_{a,2}}, \quad a = 3,4,5,6 \quad (29)$$



$$v_{2,1}(t) = \frac{NV_{scenario} \mathbf{1}(\sum_{a=3,\dots,6} \sum_{v=3,\dots,5} N_{a,v} \geq 0.8 \sum_{a=3,\dots,6} N_a)}{T_{scenario} N_{2,2}},$$

For 1st sept 2021 + 56 days  $\leq t \leq$  1st sept 2021 +  $T_{scenario}$

Uncertainty propagation in vaccination scenarios. In this study we use a mixture of full Bayesian inference, for transmission model parameters, that were specific to each Kenyan county, and minimum divergence estimators, for outcome model parameters, that were specific to Kenyan counties, e.g. the county-specific reporting/disease rate relative to Nairobi  $\psi_c$ , or were specific to particular age groups, e.g. the baseline risk of death per infection in each age group  $\mu_a^D$ .

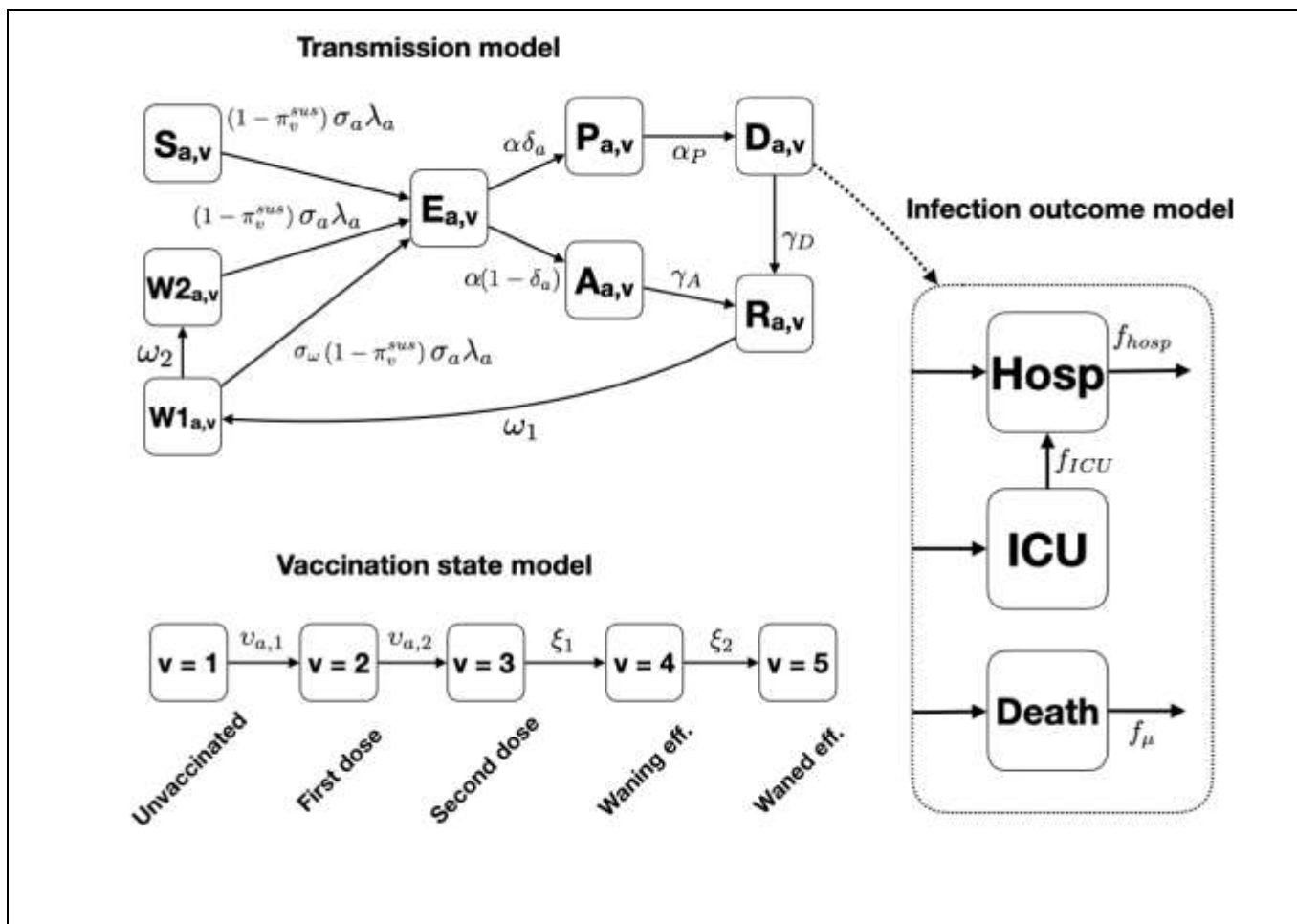
When generating scenario projections for all Kenya, we solved the transmission model (equation (1)) for each of the 47 Kenyan counties and for each of the 2000 MCMC draws of county-specific transmission model parameters. To match the 2000 MCMC draws for transmission parameters per county we drew 2000 replicates of the vaccine effectiveness from reported ranges first and second dose AstraZeneca effectiveness against Delta variant. This generated 2000 expected daily reported death incidence, ICU occupancy and general ward occupancy for each county (equations (14-15), (17) for each county) and for each of the 7 vaccine rollout scenarios.

To account for (1) uncertainty in transmission parameters, (2) unpredictability in reporting, and (3) uncertainty in vaccine effectiveness, the prediction intervals for Kenya as a total were calculated by

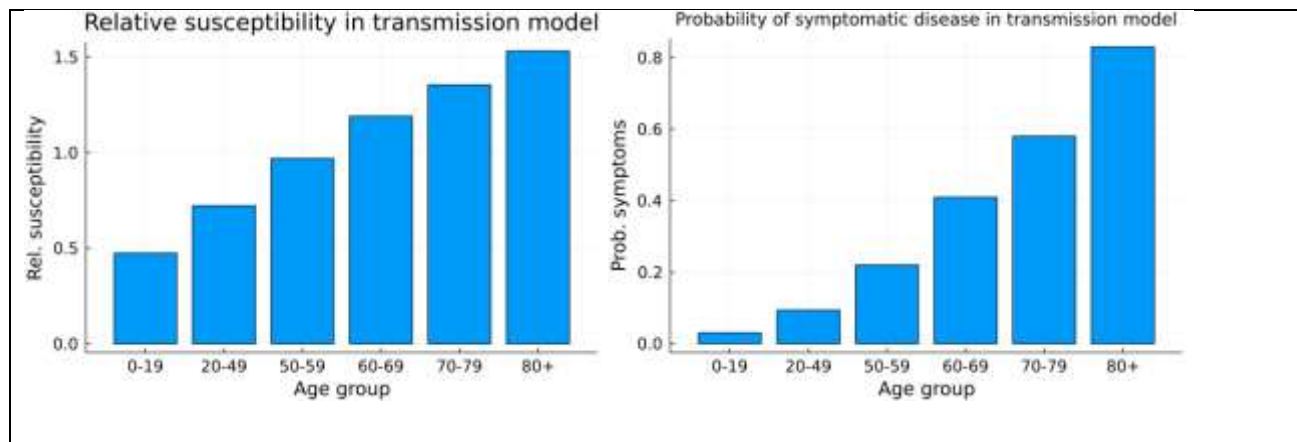
1. Augmenting the 2000 projections of expected daily reported observables per day, per age group, and per county into a single group of 2000 projections of expected daily reported observables per day, and per age group by summing across counties.
2. Converting from *expected* daily reported observables to *random* instances by sampling from the negative binomial distribution that minimized divergence between actual observed data and the model projections (equations (25-27)) for each day of the 2000 projections.
3. Presenting the daily ensemble average (and 95% ensemble prediction intervals) across the 2000 randomized projections.

Modelling immune escape variant. In this paper we consider an immune escape variant which reduces protection against reinfection by 50% across natural immunity and vaccine protection, and spreads faster through the population with a 30% reduced generation time but is otherwise epidemiologically like the Delta variant. Concretely, we implemented this by assuming that the immune escape variant arrived in Kenya on 15<sup>th</sup> November 2021, during a period of low infections for other variants, and applying a set of instantaneous effects:

- Relative immune escape variant frequency becomes 100% in all Kenyan counties, reflecting rapid dominance of invading variant.
- All transmission rate parameters (e.g.  $\beta_0, \alpha, \alpha_p, \gamma_D, \gamma_A$ ) increase by a factor 1/0.7.
- 50% of people in fully or partially immune (post-natural infection) categories (**R**, **W1**) transition instantly to completely waned immunity to reinfection (**W2**).
- Vaccine effectiveness against reinfection and reduction of infectiousness ( $\pi_2^{sus}, \pi_3^{sus}, \pi_2^{inf}, \pi_3^{inf}$ ) decrease by 50%.



**Figure S1:** Schematic diagrams of the main modelling components. **Top left:** The transmission model for each Kenyan county follows an essentially SEIR type pattern with separate categories for symptomatic and asymptomatic infection, and two categories of waning natural immunity. **Right:** The infection outcome model for a fraction of infections that are either deadly, critical (lead to a period in ICU followed by a period in general wards), or severe (lead to a period in general wards). **Bottom left:** Vaccination state model. Transitions indicate per capita inoculation rates with first and second doses of vaccine followed by two categories of waning vaccine-induced immunity.



**Fig S2:** Fixed age-dependent rates. **Left:** Relative susceptibility per infectious contact by age group. **Right:** Probability of a symptomatic episode per infection for each age group.

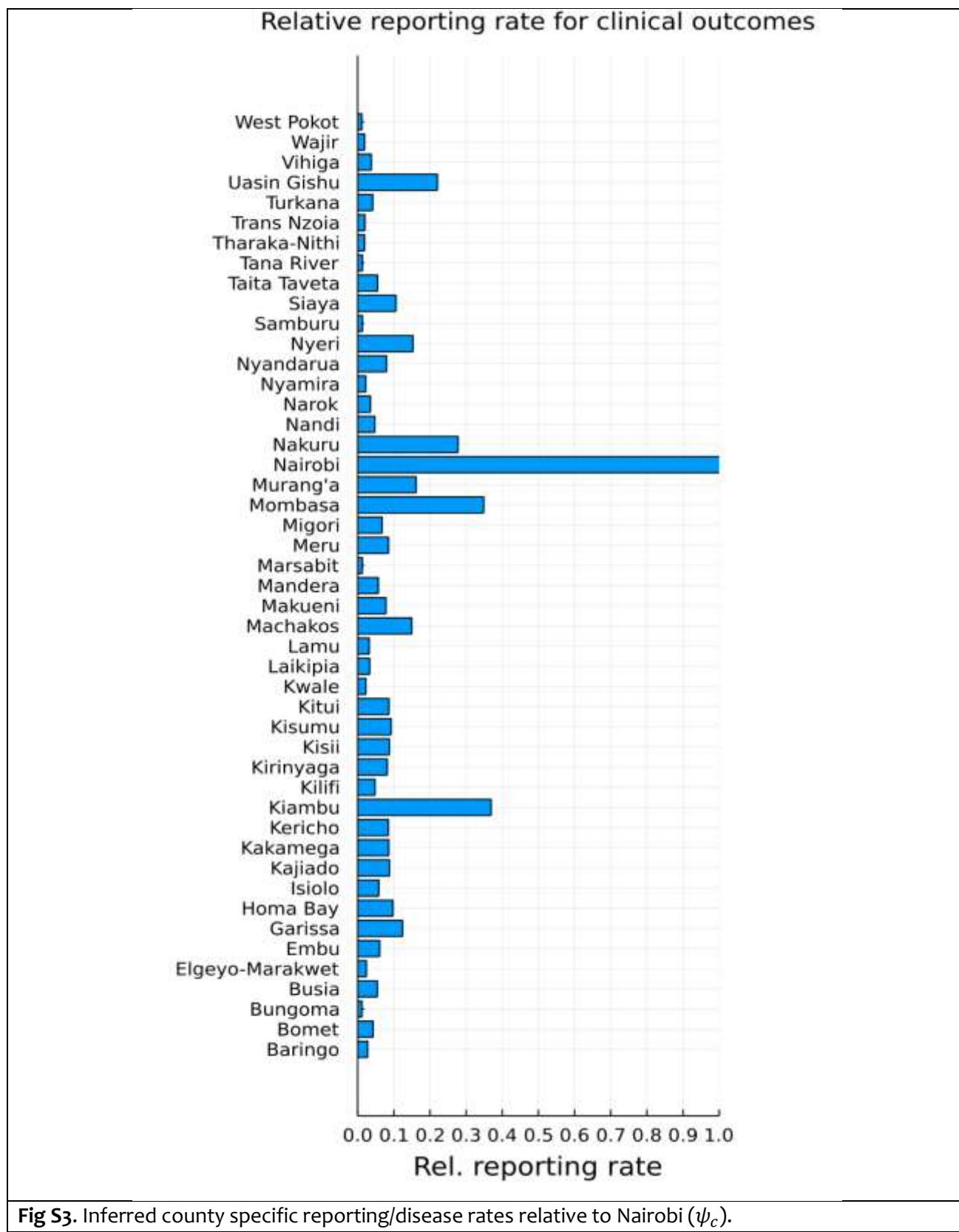


Fig S3. Inferred county specific reporting/disease rates relative to Nairobi ( $\psi_c$ ).

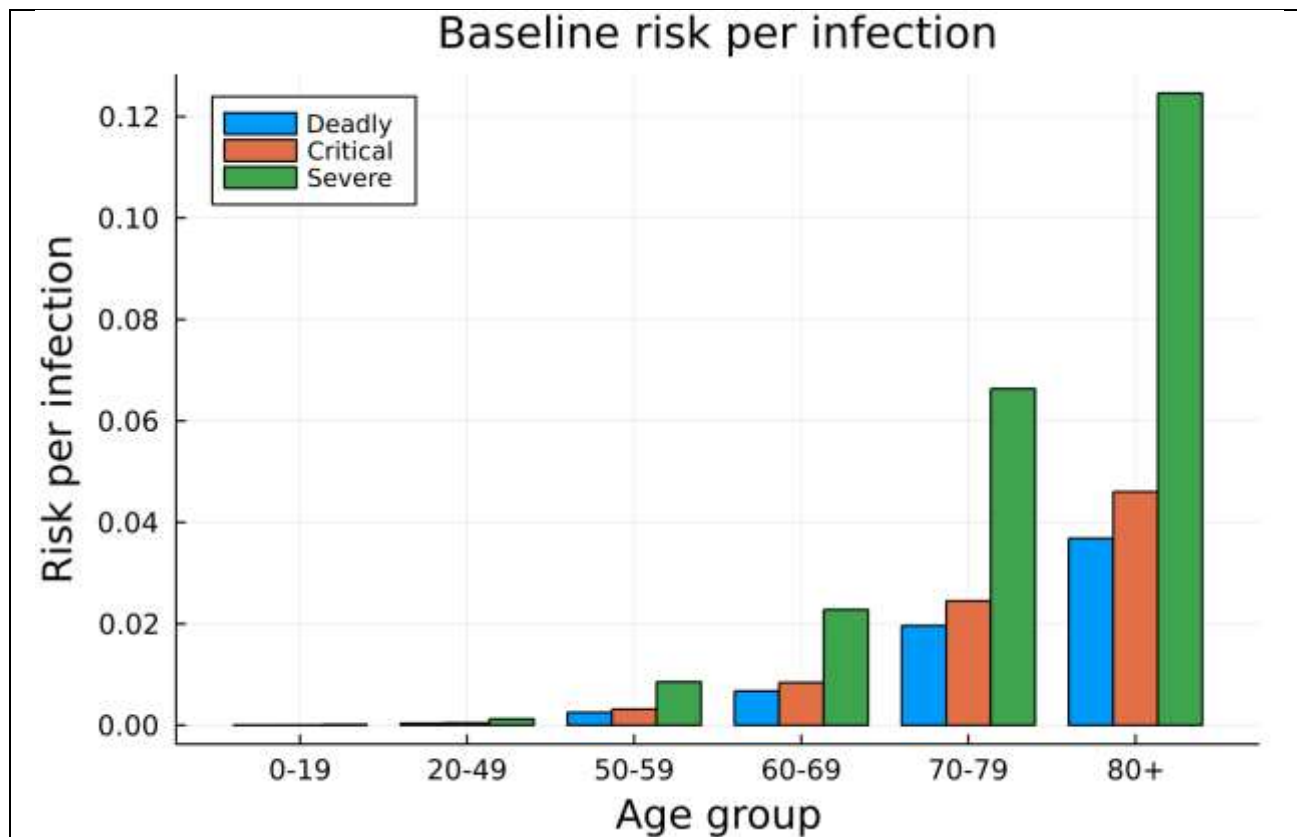


Fig S4. Inferred age-specific risks per infection in unvaccinated and naïve individuals.

**Table S1:** Transmission model parameters

Transmission model parameters and variables	Value
<b>State Variables</b>	
Number of susceptible people at time $t$ , $S(t)$	Dynamic
Number of latently infected people at time $t$ , $E(t)$	Dynamic
Number of infectious people at time $t$ , $I(t)$	Dynamic
Number of recovered and immune people at time $t$ , $R(t)$	Dynamic
Number of people who have lost natural immunity at time $t$ , $W(t)$	Dynamic
Cumulative number of primary infections at time $t$ , $(F(t))$	Dynamic
Cumulative number of all infections at time $t$ , $C(t)$	Dynamic
<b>Transmission model parameters</b>	
Number of latently infected people at time $t=0$ , 1 <sup>st</sup> January 2021, by age group, $E_a(0)$ , $a = 1,2, \dots,6$	<b>Inferred from data.</b>
Background transmission, $\beta_0$	<b>Inferred from data</b>
Scaling from social to effective contacts at home, $\beta_{home}$ , school, $\beta_{school}$ , work, $\beta_{work}$ and other places, $\beta_{other}$	<b>Inferred from data</b>
Maximum increase in transmission due to the introduction of the alpha and beta variants, $L_{\alpha\beta}$ , or the delta variant, $L_{\delta}$	<b>Inferred from data</b>
Growth rate of the alpha and beta variants, $\kappa_{\alpha\beta}$ , or the delta variant, $\kappa_{\delta}$	<b>Inferred from data</b>
Midpoint of the sigmoid growth curve for alpha and beta variants, $t_{0_{\alpha\beta}}$ , or the delta variant, $t_{0_{\delta}}$	<b>Inferred from data</b>
A factor that translates the initial (1 <sup>st</sup> Jan -10 <sup>th</sup> Mar 2021) seroprevalence by age group to initial attack rates, $\tau$	<b>Inferred from data</b>
Infectious period, $1/\gamma$	2.4 days. Chosen to recreate a serial interval of 5.5days [58]
Latent period, $1/\sigma$	3.1 days. The mean incubation period [49] was reduced by two days of pre-symptomatic transmission [59] to give a latency period.
Mean period of complete protection after a natural infection, $1/\omega$	180 days, point estimate based on reinfection studies [60–63]
Relative susceptibility compared to naïve individuals after the loss of complete protection after the first infection, $\sigma_{\omega}$	$\sigma_{\omega}=0.16$ . Point estimate based on reinfection studies [60–63]
$c_t$ – contact rate	1, Equivalent to assuming contacts are back to baseline and stable



$v_i$ - Vaccine effectiveness against transmission (delta variant)	(0% to 35.0%)--dose 1,(0% to 69.0%)--dose 2 [15]
$v_a$ - Vaccine effectiveness against acquisition (delta variant)	(55% to 65%)--dose 1,(65% to 80%)--dose 2 [14]
$v_d$ - Vaccine effectiveness against severe disease (delta variant)	(80% to 90%) dose 1,(95% to 99%)--dose 2 [14]
$v_\mu$ - Vaccine effectiveness against death (delta variant)	(90%-95%) -dose 1, (95% to 99%) -dose 2 [14]
$1/r_{vp_i}$ - rate of vaccine progression to full efficacy	14 days after each dose ( $i = 1,2$ ) [64]
<b>Observation model parameters and data</b>	
Number of people, by age group, who would test PCR positive on day $n$ , $(P^+)_n$ .	Dynamic
Number of people, by age group, who were observed to test PCR positive on day $n$ , $(P_{obs}^+)_n$ .	Data
Number of people, by age group, who would test as sero-converted on day $n$ , $(S^+)_n$ .	Dynamic
Number of people, by age group, who actually test as sero-converted on day $n$ , $(S_{obs}^+)_n$ .	Data
Probability that an infected individual would test PCR positive on day $t$ after infection, $Q_{PCR}(t)$	$Q_{PCR}(t) = f_{onset} Q_\Gamma(\tau)$ where $Q_\Gamma(\tau)$ was the tail function of a gamma distribution fitted to data given in [65] and is the probability function of onset of symptoms post-infection [49].
Probability that an infected individual would be detectably seropositive on day $t$ after infection, $Q_{sero}(t)$	$Q_{sero}(t)$ is linearly increasing over 26 days to saturate at 92.7% sensitivity, based on report delay in seroconversion [65] and maximum sensitivity of serological assay [66].
Relative bias in favour of selecting a PCR positive individual for testing, $\chi$	<b>Inferred from data</b>
<b>Clinical outcome parameters</b>	
The age-dependent probability of death given severe infection, $\mu_a^D$	<b>Inferred from data</b>
The age-dependent probability of critical disease given severe infection, $\mu_a^C$	<b>Inferred from data</b>
The age-dependent probability of severe disease given severe infection, $\mu_a^S$	<b>Inferred from data</b>
Variant specific risk of death, $\psi_a^D$	<b>Inferred from data</b>
Variant specific risk of critical disease, $\psi_a^C$	<b>Inferred from data</b>
Variant specific risk of severe disease, $\psi_a^S$	<b>Inferred from data</b>

## Economic Evaluation Equations

### Productivity losses

Productivity losses was calculated using the following equation:

$$PL = PL_{morbidity} + PL_{mortality} \quad (a)$$

Where: PL=productivity losses;  $PL_{morbidity}$ =productivity loss due to morbidity;  $PL_{mortality}$ =productivity loss due to mortality

$$PL_{morbidity} = PL_{asymptomatic} + PL_{mild} + PL_{severe} + PL_{critical}$$

$$PL_{asymptomatic} = \text{Number of asymptomatic cases} \times \text{testing rate} \times \text{GDP per capita} \times \text{quarantine period} \times \text{informal sector proportion}$$

$$PL_{mild} = \text{Number of mild cases} \times \text{testing rate} \times \text{GDP per capita} \times \text{quarantine period} \times \text{informal sector proportion}$$

$$PL_{severe} = \text{Number of severe cases} \times \text{GDP per capita} \times \text{quarantine period}$$

$$PL_{critical} = \text{Number of critical cases} \times \text{GDP per capita} \times \text{duration of critical disease}$$

The testing rate was used to apportion the number of asymptomatic and mild cases who are tested. Further the proportion of informal sector is used to apply lost productivity on asymptomatic/mild cases that are in the informal sector, given the assumption that only those in informal sector are likely not to be productive as they isolate. Lastly, the duration of disease is used where the duration of illness is more than the 14 day quarantine period.

$$PL_{mortality} = YLL \times \text{GDP per capita}$$

Where: YLL=Years of life lost (described below)

### Disability adjusted life years (DALYs)

Disability adjusted life years (DALYs) was calculated using the equation:

$$DALYs = \sum_a^j YLL_a + \sum_h^k YLD_h \quad (b)$$

Where:  $a$  is the age at death;  $j$ = number of age groups;  $h$  are the health states;  $k$  =number of health states.

YLL is estimated as [22]:

$$YLL = \frac{KCe^{ra}}{(r + \beta)^2} \{e^{-(r+\beta)(L+a)}[-(r + \beta)(L + a) - 1] - e^{-(r+\beta)a}[-(r + \beta)a - 1]\} + \frac{1 - K}{r}(1 - e^{-rL}) \quad (c)$$

Where: K=age weighting modulating factor; C=adjustment constant for age weights; r=discount rate; a=age of death;  $\beta$ =parameter from the age weighting function; L=standard life expectancy at age of death

Years lost due to disability (YLD) is estimated as [22]:

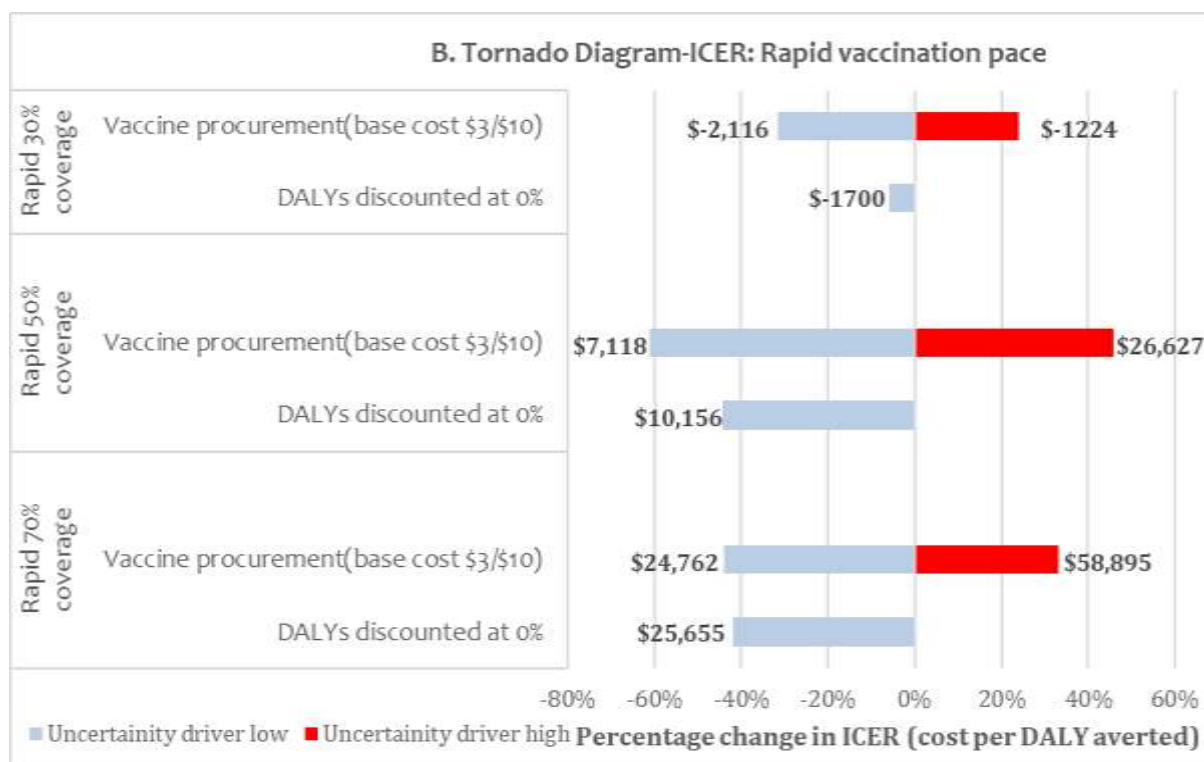
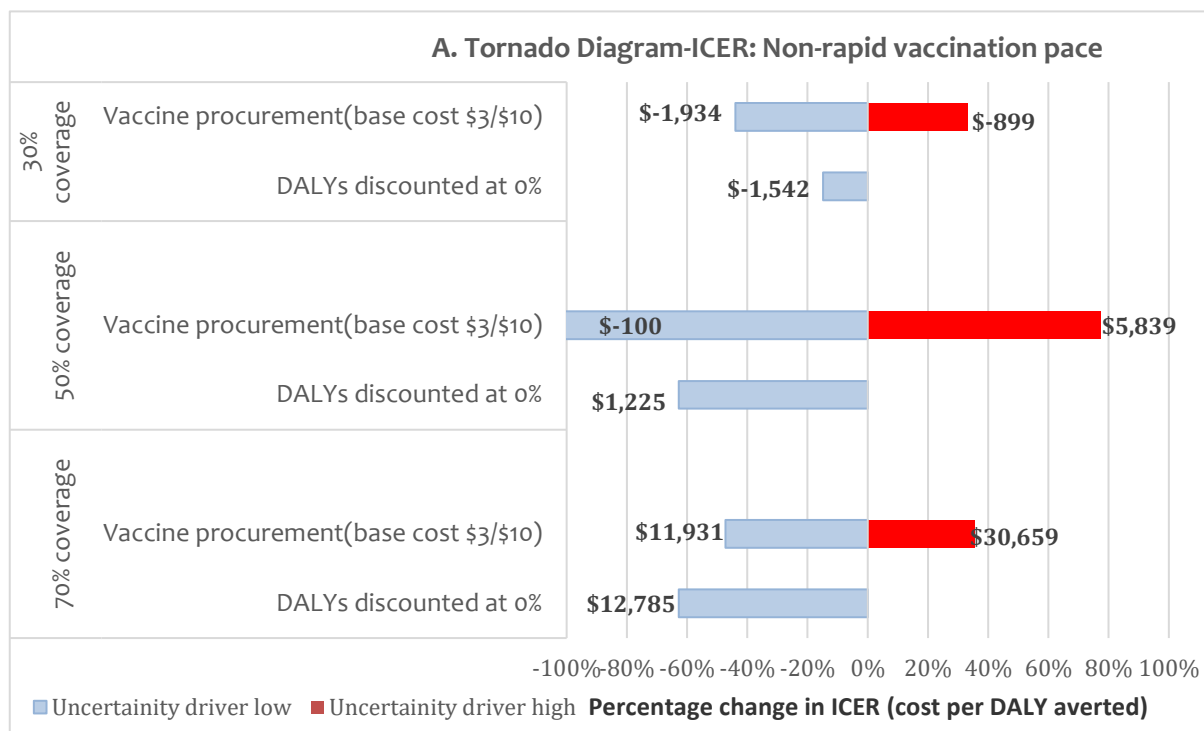
$$\text{YLDs} = D \left\{ \frac{K C e^{ra}}{(r + \beta)^2} \left\{ e^{-(r+\beta)(L+a)} [-(r + \beta)(L + a) - 1] - e^{-(r+\beta)a} [-(r + \beta)a - 1] \right\} + \frac{1 - K}{r} (1 - e^{-rL}) \right\} \quad (\text{d})$$

Where: K=age weighting modulating factor; C=adjustment constant for age weights; r=discount rate; a=age of onset of disability;  $\beta$ =parameter from the age weighting function; L=duration of disability; D=disability weight

**Table S2:** Statistical distributions for the probabilistic sensitivity analysis

	<b>Statistical distribution*</b>
<b>Costs</b>	
Vaccine delivery cost per patient at 30% coverage	Gamma (10055,6.1E-4)
Vaccine delivery cost per patient at 50% coverage	Gamma (13679,3E-4)
Vaccine delivery cost per patient at 70% coverage	Gamma (9221,4E-4)
Unit treatment cost of an asymptomatic patient	Gamma (103.1,9.6E-4)
Unit treatment cost of a mild case	Gamma (105.1,1.8E-1)
Unit treatment cost of a severe case	Gamma (105.1,1.2)
Unit treatment cost of a critical case	Gamma (105.1,5.7)
<b>Health outcomes</b>	
Disability weight for a critical case	Beta (102.8,54.2)
Disability weight for a severe case	Beta (21.6,140.8)
Disability weight for a mild case	Beta (21.3,396.2)
Length of hospitalization for severe episode	Gamma (53.3,3.6E-4)
Length of ICU stay for critical episode	Gamma (53.3,3.6E-4)

\*The form of the distributions was chosen according to the nature of the parameter. (i.e. parameters describing the probability of health outcomes occurrence of an event are given by a beta distribution, while the values of the vaccine delivery costs and length of hospitalization/ICU stay of the various scenarios were described using a gamma distribution).



**Figure S6:** One-way sensitivity analysis of vaccine prices and discounting rate on ICERs

**Table S3:** Projected costs and the cost-effectiveness of different vaccination strategies in Kenya from a health system’s perspective

	Economic Outcomes				
	Total vaccine costs (\$ million) Median (2.5 <sup>th</sup> to 97.5 <sup>th</sup> percentile)	Total treatment costs (\$million) Median (2.5 <sup>th</sup> to 97.5 <sup>th</sup> percentile)	Total health care costs (\$million) Median (2.5 <sup>th</sup> to 97.5 <sup>th</sup> percentile)	Total DALYs (thousands) Median(95%CI)	ICER, USD per DALY averted Mean (95%CI)
<b>Non-rapid vaccination pace (administered within 1.5years)</b>					
No vaccination	-	313 (269 to 405)	313 (269 to 405)	247 (243 to 252)	-
30% coverage	236 (233 to 238)	157 (135 to 199)	393 (371 to 436)	114 (110 to 118)	555 (553 to 557)
50% coverage	323 (321 to 324)	140 (120 to 177)	463 (444 to 501)	101 (97 to 104)	5,195 (5,190 to 5,199)
70% coverage	443 (441 to 445)	133 (115 to 169)	576 (560 to 612)	96 (92 to 100)	24,535 (24,514 to 24,557)
<b>Rapid vaccination pace (administered within 6 months)</b>					
No vaccination	-	313 (269 to 405)	313 (269 to 405)	247 (243 to 252)	-
30% coverage	236 (233to 238)	129 (111 to 163)	366 (347 to 399)	93 (89 to 96)	291 (290 to 295)
50% coverage	323 (321 to 324)	124 (107 to 157 )	447 (430 to 481)	88 (85 to 92)	20,157 (20,126 to 20,128)
70% coverage	443 (441 to 445)	121 (104 to 152)	564 (549 to 596)	86 (82 to 89)	46,150 (46,025 to 46,274)

Costs=rounded off to the nearest 1,000,000; Total DALY rounded off to the nearest 1,000; ICERs=rounded off to the nearest whole number



## REFERENCES

- 1 Mathieu E, Ritchie H, Ortiz-Ospina E, et al. Coronavirus Pandemic (COVID-19). *Our World Data* Published Online First: 2021.<https://ourworldindata.org/coronavirus> (accessed 15 Dec 2021).
- 2 World Health Organization. COVID-19 vaccines. 2021.<https://www.who.int/emergencies/diseases/novel-coronavirus-2019/covid-19-vaccines> (accessed 20 Oct 2021).
- 3 Taylor L. Covid-19: Hong Kong reports world's highest death rate as zero covid strategy fails. *BMJ* 2022;**376**:o707. doi:10.1136/BMJ.O707
- 4 Brand SPC, Ojal J, Aziza R, et al. COVID-19 transmission dynamics underlying epidemic waves in Kenya. *Science (80- )* Published Online First: 7 October 2021. doi:10.1126/science.abk0414
- 5 Chisale MRO, Ramazanu S, Mwale ES, et al. Seroprevalence of anti-SARS-CoV-2 antibodies in Africa: A systematic review and meta-analysis. *Rev Med Virol* Published Online First: 6 July 2021. doi:10.1002/RMV.2271
- 6 Ministry of Health Kenya. National COVID-19 Vaccine Deployment Plan. 2021.
- 7 Ministry of Health. Kenya COVID-19 Vaccination Program-Daily Situation Report. 2022. <https://www.health.go.ke/wp-content/uploads/2022/03/MINISTRY-OF-HEALTH-KENYA-COVID-19-IMMUNIZATION-STATUS-REPORT-20TH-MARCH-2022.pdf> (accessed 22 Mar 2022).
- 8 Moore S, Hill EM, Dyson L, et al. Modelling optimal vaccination strategy for SARS-CoV-2 in the UK. *PLOS Comput Biol* 2021;**17**:e1008849. doi:10.1371/JOURNAL.PCBI.1008849
- 9 Moore S, Hill EM, Tildesley MJ, et al. Vaccination and non-pharmaceutical interventions for COVID-19: a mathematical modelling study. *Lancet Infect Dis* 2021;**21**:793–802. doi:10.1016/S1473-3099(21)00143-2/ATTACHMENT/6721682C-734C-4335-A43B-6F5B618390E1/MMC1.PDF
- 10 Uyoga S, Adetifa IMO, Otiende M, et al. Prevalence of SARS-CoV-2 Antibodies From a National Serosurveillance of Kenyan Blood Donors, January-March 2021. *JAMA* 2021;**326**:1436–8. doi:10.1001/JAMA.2021.15265
- 11 Etyang AO, Adetifa I, Omoro R, et al. SARS-CoV-2 Seroprevalence in three Kenyan Health and Demographic Surveillance Sites, December 2020-May 2021. *medRxiv* Published Online First: 9 February 2022. doi:10.1101/2022.02.07.22270012
- 12 Rees EM, Nightingale ES, Jafari Y, et al. COVID-19 length of hospital stay: a systematic review and data synthesis. *BMC Med* 2020 **18**:1–22. doi:10.1186/S12916-020-01726-3
- 13 Kenya National Bureau of Statistics. 2019 Kenya Population and Housing Census: Volume III. 2019.
- 14 UK Health Security Agency. COVID-19 vaccine surveillance report: Week 11. 2022. [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/1061532/Vaccine\\_surveillance\\_report\\_-\\_week\\_11.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1061532/Vaccine_surveillance_report_-_week_11.pdf) (accessed 6 Apr 2022).
- 15 Clifford S, Waight P, Hackman J, et al. Effectiveness of BNT162b2 and ChAdOx1 against SARS-CoV-2 household transmission: a prospective cohort study in England. *medRxiv* Published Online First: 25 November 2021. doi:10.1101/2021.11.24.21266401
- 16 Keeling MJ, Thomas A, Hill EM, et al. Waning, Boosting and a Path to Endemicity for SARS-CoV-2. *medRxiv* Published Online First: 10 November 2021. doi:10.1101/2021.11.05.21265977
- 17 Drummond M, Sculpher M, Claxton K, et al. *Methods for the economic evaluation of health care programmes*. Fourth Edition. Oxford: : Oxford University Press 2015.
- 18 Johns B, Baltussen R, Hutubessy R. Programme costs in the economic evaluation of health interventions. *Cost Eff Resour Alloc* 2003;**1**:1–10. doi:10.1186/1478-7547-1-1/TABLES/3
- 19 Orangi S, Kairu A, Ngatia A, et al. Examining the unit costs of COVID-19 vaccine delivery in Kenya. *BMC Heal Serv Res* 2022;**22**. doi:10.1186/S12913-022-07864-Z

- 20 Barasa E, Kairu A, Ng'ang'a W, et al. Examining unit costs for COVID-19 case management in Kenya. *BMJ Glob Heal* 2021;**6**. doi:10.1136/BMJGH-2020-004159
- 21 Pritchard C, Sculpher M. Productivity costs: Principles and practice in economic evaluation. 2000.
- 22 Fox-Rushby JA, Hanson K. Calculating and presenting disability adjusted life years (DALYs) in cost-effectiveness analysis. *Health Policy Plan* 2001;**16**:326–31. doi:10.1093/HEAPOL/16.3.326
- 23 World Health Organization. Life tables: Kenya. Glob. Heal. Obs. data Repos. 2020.<https://apps.who.int/gho/data/view.main.60850?lang=en> (accessed 20 Oct 2021).
- 24 Luangsanatip N, Prawjaeng J, Saralamba S, et al. Optimal vaccine strategy to control COVID-19 pandemic in middle-income countries: Modelling case study of Thailand. *Res Sq Published Online First*: 24 February 2021. doi:10.21203/RS.3.RS-270635/V1
- 25 Salomon JA, Haagsma JA, Davis A, et al. Disability weights for the Global Burden of Disease 2013 study. *Lancet Glob Heal* 2015;**3**:e712–23. doi:10.1016/S2214-109X(15)00069-8
- 26 Haagsma JA, Maertens De Noordhout C, Polinder S, et al. Assessing disability weights based on the responses of 30,660 people from four European countries. *Popul Health Metr* 2015;**13**. doi:10.1186/s12963-015-0042-4
- 27 Ochalek J, Lomas J, Claxton K. Estimating health opportunity costs in low-income and middle-income countries: a novel approach and evidence from cross-country data. *BMJ Glob Heal* 2018;**3**:e000964. doi:10.1136/BMJGH-2018-000964
- 28 Woods B, Revill P, Sculpher M, et al. Country-Level Cost-Effectiveness Thresholds: Initial Estimates and the Need for Further Research. *Value Heal* 2016;**19**:929–35. doi:10.1016/J.JVAL.2016.02.017
- 29 World Health Organization. Life tables by country - Kenya. Glob. Heal. Obs. data Repos. 2020.
- 30 World Bank. World Bank Open Data . 2021.<https://data.worldbank.org/> (accessed 11 Nov 2021).
- 31 Voysey M, Costa Clemens SA, Madhi SA, et al. Single-dose administration and the influence of the timing of the booster dose on immunogenicity and efficacy of ChAdOx1 nCoV-19 (AZD1222) vaccine: a pooled analysis of four randomised trials. *Lancet* 2021;**397**:881–91. doi:[https://doi.org/10.1016/S0140-6736\(21\)00432-3](https://doi.org/10.1016/S0140-6736(21)00432-3)
- 32 Eyre DW, Taylor D, Purver M, et al. The impact of SARS-CoV-2 vaccination on Alpha & Delta variant transmission. *medRxiv Published Online First*: 29 September 2021. doi:10.1101/2021.09.28.21264260
- 33 Ioannou GN, Locke ER, O'Hare AM, et al. COVID-19 Vaccination Effectiveness Against Infection or Death in a National U.S. Health Care System. *Ann Intern Med* Published Online First: 21 December 2021. doi:10.7326/M21-3256
- 34 Sheikh A, Robertson C, Taylor B. BNT162b2 and ChAdOx1 nCoV-19 Vaccine Effectiveness against Death from the Delta Variant. *N Engl J Med* 2021;**385**:2195–7. doi:10.1056/NEJMC2113864/SUPPL\_FILE/NEJMC2113864\_DISCLOSURES.PDF
- 35 Tembo J, Maluzi K, Egbe F, et al. Covid-19 in Africa. *BMJ* 2021;**372**. doi:10.1136/BMJ.N457
- 36 Saltelli A, Annoni P, Azzini I, et al. Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index. *Comput Phys Commun* 2010;**181**:259–70. doi:10.1016/J.CPC.2009.09.018
- 37 Reddy KP, Fitzmaurice KP, Scott JA, et al. Clinical outcomes and cost-effectiveness of COVID-19 vaccination in South Africa. *Nat Commun* 2021;**12**:6238. doi:10.1038/S41467-021-26557-5
- 38 Moghadas SM, Vilches TN, Zhang K, et al. The Impact of Vaccination on Coronavirus Disease 2019 (COVID-19) Outbreaks in the United States. *Clin Infect Dis* 2021;**73**:2257–64. doi:10.1093/CID/CIAB079
- 39 Pearson CAB, Bozzani F, Procter SR, et al. COVID-19 vaccination in Sindh Province, Pakistan: A modelling study of health impact and cost-effectiveness. *PLoS Med* 2021;**18**.

- doi:10.1371/JOURNAL.PMED.1003815
- 40 Rasambainarivo F, Ramiadantsoa T, Raherinandrasana A, et al. Prioritizing COVID-19 vaccination efforts and dose allocation within Madagascar. *BMC Public Heal* 2022;**22**. doi:10.1186/S12889-022-13150-8
- 41 Hagens A, İnkaya AC, Yildirak K, et al. COVID-19 Vaccination Scenarios: A Cost-Effectiveness Analysis for Turkey. *Vaccines* 2021;**9**. doi:10.3390/VACCINES9040399
- 42 Gessner BD, Kaslow D, Louis J, et al. Estimating the full public health value of vaccination. *Vaccine* 2017;**35**:6255–63. doi:10.1016/J.VACCINE.2017.09.048
- 43 Wilder-Smith A, Longini I, Zuber PL, et al. The public health value of vaccines beyond efficacy: Methods, measures and outcomes. *BMC Med* 2017;**15**:1–9. doi:10.1186/S12916-017-0911-8/TABLES/2
- 44 Rodrigues CMC, Plotkin SA. Impact of Vaccines; Health, Economic and Social Perspectives. *Front Microbiol* 2020;**11**:1526. doi:10.3389/FMICB.2020.01526/BIBTEX
- 45 Prem K, van Zandvoort K, Klepac P, et al. Projecting contact matrices in 177 geographical regions: An update and comparison with empirical data for the COVID-19 era. *PLOS Comput Biol* 2021;**17**:e1009098. doi:10.1371/JOURNAL.PCBI.1009098
- 46 Google L.L.C. COVID-19 Community Mobility Reports. 2022.<https://www.google.com/covid19/mobility/> (accessed 4 Mar 2022).
- 47 Freunde von GISAID e.V. GISAID-Tracking of Variants. 2022.<https://www.gisaid.org/hcov19-variants/> (accessed 4 Mar 2022).
- 48 Githinji G, de Laurent ZR, Mohammed KS, et al. Tracking the introduction and spread of SARS-CoV-2 in coastal Kenya. *Nat Commun* 2021 121 2021;**12**:1–10. doi:10.1038/s41467-021-25137-x
- 49 Lauer SA, Grantz KH, Bi Q, et al. The Incubation Period of Coronavirus Disease 2019 (COVID-19) From Publicly Reported Confirmed Cases: Estimation and Application. *Ann Intern Med* 2020;**172**:577–82. doi:10.7326/M20-0504
- 50 Moghadas SM, Shoukat A, Fitzpatrick MC, et al. Projecting hospital utilization during the COVID-19 outbreaks in the United States. *Proc Natl Acad Sci U S A* 2020;**117**:9122–6. doi:10.1073/PNAS.2004064117/SUPPL\_FILE/PNAS.2004064117.SAPP.PDF
- 51 Bakilana A, Demombynes G, Gubbins P. Could COVID-19 deaths in Kenya be higher than official records? 2022. <https://blogs.worldbank.org/african/could-covid-19-deaths-kenya-be-higher-official-records> (accessed 10 Mar 2022).
- 52 Keeling MJ, Hill EM, Gorsich EE, et al. Predictions of COVID-19 dynamics in the UK: Short-term forecasting and analysis of potential exit strategies. *PLOS Comput Biol* 2021;**17**:e1008619. doi:10.1371/JOURNAL.PCBI.1008619
- 53 Van Wees J-D, Osinga S, Van Der Kuip M, et al. Forecasting hospitalization and ICU rates of the COVID-19 outbreak: an efficient SEIR model. *World Heal Organ Bull* Published Online First: 2020. doi:10.2471/BLT.20.251561
- 54 Rackauckas C, Nie Q. DifferentialEquations.jl – A Performant and Feature-Rich Ecosystem for Solving Differential Equations in Julia. *J Open Res Softw* 2017;**5**:15. doi:10.5334/JORS.151
- 55 Ngere I, Dawa J, Hunsperger E, et al. High seroprevalence of SARS-CoV-2 but low infection fatality ratio eight months after introduction in Nairobi, Kenya. *Int J Infect Dis* 2021;**112**:25–34. doi:10.1016/J.IJID.2021.08.062
- 56 Neal RM. MCMC Using Hamiltonian Dynamics. In: *Handbook of Markov Chain Monte Carlo*. CRC Press 2011.
- 57 Hoffman MD, Gelman A. The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo. *J Mach Learn Res* 2014;**15**:1351–81.<http://mcmc-jags.sourceforge.net> (accessed 10 Mar 2022).
- 58 Ferretti L, Wymant C, Kendall M, et al. Quantifying SARS-CoV-2 transmission suggests

- epidemic control with digital contact tracing. *Science* (80-) 2020;**368**.  
doi:10.1126/science.abb6936
- 59 Tindale LC, Stockdale JE, Coombe M, et al. Evidence for transmission of COVID-19 prior to symptom onset. *Elife* 2020;**9**:1–34. doi:10.7554/ELIFE.57149
- 60 Harvey RA, Rassen JA, Kabelac CA, et al. Association of SARS-CoV-2 Seropositive Antibody Test With Risk of Future Infection. *JAMA Intern Med* 2021;**181**:672–9. doi:10.1001/jamainternmed.2021.0366
- 61 Hall VJ, Foulkes S, Charlett A, et al. SARS-CoV-2 infection rates of antibody-positive compared with antibody-negative health-care workers in England: a large, multicentre, prospective cohort study (SIREN). *Lancet* 2021;**397**:1459–69. doi:10.1016/S0140-6736(21)00675-9
- 62 Vitale J, Mumoli N, Clerici P, et al. Assessment of SARS-CoV-2 Reinfection 1 Year After Primary Infection in a Population in Lombardy, Italy. *JAMA Intern Med* 2021;**181**:1407–8. doi:10.1001/jamainternmed.2021.2959
- 63 Lumley SF, Wei J, O'Donnell D, et al. The Duration, Dynamics, and Determinants of Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) Antibody Responses in Individual Healthcare Workers. *Clin Infect Dis* 2021;**73**:e699–709. doi:10.1093/cid/ciab004
- 64 Ewer KJ, Barrett JR, Belij-Rammerstorfer S, et al. T cell and antibody responses induced by a single dose of ChAdOx1 nCoV-19 (AZD1222) vaccine in a phase 1/2 clinical trial. *Nat Med* 2020;**27**:270–8. doi:10.1038/s41591-020-01194-5
- 65 Lisboa Bastos M, Tavaziva G, Abidi SK, et al. Diagnostic accuracy of serological tests for covid-19: systematic review and meta-analysis. *BMJ* 2020;**370**:2516. doi:10.1136/bmj.m2516
- 66 Uyoga S, Adetifa IMO, Karanja HK, et al. Seroprevalence of anti-SARS-CoV-2 IgG antibodies in Kenyan blood donors. *Science* (80-) 2021;**371**:79–82. doi:10.1126/science.abe1916