Modelling COVID-19 spread to reconstruct the viral dissemination and inform the design of interventions

Benjamin Faucher

SORBONNE UNIVERSITÉ

THÈSE DE DOCTORAT

Modélisation de la pandémie de COVID-19 pour reconstruire la dissémination du virus et informer la mise en place d'interventions.

Auteur:
Benjamin FAUCHER

Encadrants:
Pierre-Yves BOËLLE,
Chiara POLETTO

Laboratoire:
Equipe SUMO, IPLESP UMR-S 1136, INSERM

Devant un jury composé de:

Forence Débarre	Examinatrice et Présidente du jury
Pascal Crépey	Examinateur
Harold Noël	Membre du jury
Pierre-Yves Boëlle	Directeur de thèse
Chiara Poletto	Encadrante

Thèse soumise pour le grade de Docteur en Biostatistiques et Biomathématiques

Ecole Doctorale Pierre Louis de Santé Publique:
Epidémiologie et Sciences de l'Information Biomédicale

Abstract

L'émergence de nouveaux pathogènes est une préoccupation importante pour les autorités de santé publique. Dans le contexte de la pandémie de COVID-19, le SARS-CoV-2 et ses variants successifs ont suivi le même schéma. Un nouveau virus émerge dans un pays, se dissémine à l'échelle internationale, et crée une augmentation rapide des cas à travers le monde. Pour faire face à cette situation, il est important de surveiller l'épidémie, déchiffrer les données de surveillance incomplètes ou incohérentes et de rapidement mettre en place des interventions. Les modèles mathématiques peuvent aider à interpréter cela. Dans cette thèse, nous nous intéressons aux deux aspects. Dans un premier temps, nous avons développé un modèle mathématique pour comprendre la surveillance et les facteurs de la propagation épidémique qui concourent aux données observées. Nous avons reconstruit rétrospectivement la dissémination du variant Alpha durant l'automne 2020 à partir de données de séquençage et de trafic aérien. Dans un deuxième travail, nous nous sommes concentrés sur la partie intervention. Nous avons utilisé un modèle agent pour quantifier l'impact épidémiologique d'une vaccination réactive ciblant entreprises et écoles où les cas sont détectés. Nous avons testé la capacité de cette stratégie à mitiger l'augmentation générale des cas et à limiter la propagation d'un nouveau variant.

Remerciements

Je voudrais remercier les différentes personnes qui ont contribué au succès des trois années qui ont conduit à l'écriture de cette thèse.
Je veux d'abord remercier les deux rapporteurs de cette thèse, ainsi que les membres du jury, pour avoir pris le temps de lire le manuscrit, de m'avoir apporté des commentaires et des remarques à la fois pointues et pertinentes. Enfin, Je suis arrivé à Paris au début de 2021, en plein milieu de la pandémie de COVID, pour commencer un stage dans le laboratoire où j'ai ensuite effectué ma thèse. En regardant en arrière, je reste stupéfait de la chance que j'ai eue de rencontrer des personnes exceptionnelles tout au long de ces trois années. Ce sont ces personnes que je voudrais sincèrement remercier.
Mes premières pensées vont d'abord à mes deux encadrants qui m'ont chacun beaucoup apporté. D'abord à Chiara, qui, par ses qualités scientifiques et humaines, m'a permis d'avancer dans cette thèse et de grandir professionnellement. Par-dessus tout, sa précision tant dans l'analyse des données que dans l'écriture, mais aussi sa capacité à trouver le fragile équilibre entre supervision et autonomie, fait d'elle pour moi un exemple pour la suite de ma carrière. Ensuite, à Pierre-Yves, mon directeur de thèse, qui a toujours le bon conseil qui débloque la situation, quelle que soit la difficulté.
Je voudrais également remercier le reste de l'équipe SUMO de l'IPLESP. D'abord Vittoria Colizza, Raphaëlle Metras et Eugenio Valdano, pour l'esprit qu'ils impulsent dans cette équipe formidable où se mêlent bonne ambiance et exigence. Je voudrais ensuite remercier mes collègues, qu'ils soient ou non encore dans le laboratoire : Laura, Giulia(s), Elisabetta, Chiara, Mattia, Wen, Francesco, Marika. J'ai une pensée particulière pour le « cantine club » : Davide, Anthony, Pourya et Francesco.
Je voudrais également remercier les personnes que j'ai rencontrées en dehors du laboratoire, qui m'ont permis d'avoir une vie équilibrée et m'ont également aidé à me développer ces trois dernières années. D'abord Aymeric, Adrien et FX, qui, en plus d'être devenus des amis très proches, sont pour moi des exemples par leurs qualités humaines parmi lesquelles honnêteté, gentillesse et intelligence. Mais aussi Brice, Massy, Eva, Léna, Antoine, Basile, Victor et Hélène. Enfin, je voudrais remercier mes parents et mes frères pour leur soutien au quotidien.

Introduction

La pandémie de Covid-19 s'est caractérisée par des vagues successives de cas. Qu'il s'agisse de la souche historique apparue en Chine ou des variants ultérieurs, le processus de dissémination a été le même : un nouveau virus/variant émerge dans un pays, se propage à l'échelle mondiale et déclenche une augmentation rapide du nombre de cas dans le monde entier. À chaque étape de la dissémination, il est essentiel de surveiller l'évolution de l'épidémie afin d'obtenir les informations nécessaires pour planifier la réponse appropriée. Les modèles mathématiques sont des outils essentiels à cet égard. Ils peuvent d'abord aider à interpréter les données provenant de systèmes de surveillance hétérogènes. Ils permettent de reconstituer la propagation du virus en estimant des variables non accessibles directement. Toutefois, il ne suffit pas de comprendre la situation épidémique. Les modèles peuvent être utilisés pour simuler différentes interventions in silico. Cela permet de tester différents scénarios et de quantifier le rapport coût-bénéfice de ces interventions. Les décideurs politiques peuvent ensuite utiliser les résultats pour mettre en œuvre des mesures de santé publique.

Depuis le début de la pandémie de COVID-19, de nombreux cas n'ont pas été détectés par les systèmes de surveillance pour plusieurs raisons. Les symptômes cliniques étaient aspécifiques et pouvaient facilement être confondus avec ceux d'autres maladies respiratoires telles que la grippe. De nombreux cas sont restés totalement asymptomatiques. En outre, la progression de la maladie comportait une phase présymptomatique au cours de laquelle les individus étaient infectieux mais ne présentaient aucun symptôme, ce qui retardait la détection. Pour des raisons pratiques, il était difficile d'assurer une surveillance virologique précise. Au début de la pandémie, la détection des cas reposait uniquement sur les symptômes cliniques jusqu'à ce que les premiers tests PCR soient disponibles à la mi-janvier 2020. Si la capacité de détection des cas de COVID-19 s'est accrue tout au long de l'année 2020, l'émergence de nouveaux variants a posé de nouveaux défis, car leur détection et leur suivi ont nécessité la mise en œuvre de systèmes de séquençage. Ces systèmes exigent des ressources et des infrastructures moins disponibles dans les pays à faible revenu, ce qui a entraîné de fortes hétérogénéités dans la surveillance virologique mondiale.

Dans un premier temps, les interdictions de voyager visaient à contenir le virus à sa source et à empêcher son importation. En janvier 2020, tous les voyages à destination et en provenance de Wuhan ont été interdits , et une approche similaire a été utilisée avec l'émergence de nouveaux variants. De nombreux pays ont restreint les voyages en provenance du Royaume-Uni, de l'Afrique du Sud et de l'Inde, où les variants Alpha, Beta et Delta sont apparues . Des mesures telles que la surveillance active ou la recherche des contacts ont tenté de prévenir les cas importés de SRAS-CoV-2 et de variants successifs. Lorsque cela a été possible, la vaccination réactive a également été utilisée pour empêcher la propagation du variant Delta. Cependant, en raison de la sous-détection des premiers cas, le virus (ou les nouveaux variants) s'est propagé silencieusement et a déclenché des épidémies locales dans d'autres pays. Les autorités étaient ensuites contraintes de mettre en œuvre des mesures d'atténuation

conçues pour freiner la transmission locale : confinement, couvre-feux, distanciation sociale, recherche des contacts et, lorsqu'elle est devenue possible, vaccination. Dans cette thèse, j'ai réalisé un travail de modélisation portant sur les deux faces de la médaille : la surveillance et l'intervention. Il est tout d'abord nécessaire de comprendre la situation épidémique pour réagir correctement à la propagation d'un nouveau virus ou d'un nouveau variant. Pour cela, il faut être capable d'interpréter correctement les données de surveillance.

La pandémie de COVID-19 s'est accompagnée d'une intense production scientifique , dans laquelle la modélisation a joué un rôle important. La modélisation a été utilisée pour l'évaluation de l'épidémie, l'estimation de paramètres épidémiologiques clés (nombre de reproduction, temps de génération, part des asymptomatiques...), la compréhension de l'épidémie en temps réel, la prévision de son déroulement à plus long terme, et l'évaluation de l'efficacité des stratégies d'intervention. En fonction de l'objectif spécifique, les études de modélisation peuvent être prospectives ou rétrospectives, théoriques ou fondées sur des données. Pendant une pandémie, les résultats de la modélisation doivent être concis et exploitables, adaptés aux besoins spécifiques des décideurs. En outre, les modèles doivent être adaptés aux données disponibles pour la paramétrisation et la validation . Les efforts de modélisation déployés pendant la pandémie de COVID-19 ont utilisé des cadres élaborés pour des épidémies antérieures et adaptés à la situation actuelle. Les modèles de métapopulation ont été largement utilisés pour étudier la propagation d'un nouveau virus ou d'un variant. Ces modèles décrivent la propagation du virus à travers des sous-populations interconnectées dont l'échelle peut aller de celle d'une ville à celle d'un pays ou d'un continent. Le type de connexions (trafic aérien, trafic routier, etc.) dépend du problème spécifique et du modèle utilisé. L'utilisation de modèle compartimentaux et de modèles agent ont également largement contribué à évaluer l'efficacité de différentes mesures.

Dans le premier article présenté dans cette thèse [1], nous avons étudié la propagation internationale du variant Alpha à partir du deuxième semestre de 2020. Nous avons développé un modèle de propagation internationale utilisant des données de trafic aérien auquel nous avons ajouté une partie sur la détection. Pour cela, nous avons utilisé des métadonnées des séquences soumises sur la plateforme GISAID ([2]) (pays de séquençage, date de collecte, date de soumission à GISAID). En ajustant les dates de premières soumissions dans différents pays ainsi que la date de collecte associée, nous avons estimé la date de première introduction du variant, et donc le temps de circulation silencieuse. Nous avons pu montrer que les mesures de restriction du trafic international étaient peu efficaces pour empêcher le variant de se propager car mises en place trop tard. Néanmoins, des mesures de mitigation locales peuvent freiner l'importation du variant.

Dans le deuxième article présenté dans cette thèse ([3]), nous nous sommes concentrés sur la partie intervention. Nous avons utilisé un modèle agent pour étudier la faisabilité d'implémenter une stratégie de vaccination réactive face à l'arrivée d'un nouveau variant.

Notre étude est calibrée sur une ville française de taille moyenne (Metz) pendant la période d'arrivée du variant Delta au début de l'été. La population synthétique est modélisée par un réseau multicouche dynamique qui représente les contacts entre individus dans cinq lieux différents (lieux de travail, transport, communauté, école et foyers) et prend en compte les données réelles fournies par l'Insee (nombre et taille des foyers, des écoles, des entreprises. . .), le nombre de contacts moyen entre individus

([4]) et l'impact des mesures de distanciation sociale ([5]). Nous avons également introduit les échelles de temps pertinentes : temps pour les symptômes de se déclarer, temps pour détecter un cas, délai pour détecter les contacts, pour acheminer les vaccins... Nous avons modélisé l'état des individus par un modèle compartimental qui prend en compte la vaccination.

Nous avons choisi d'étudier ici une vaccination réactive qui cible les lieux de travail et les écoles. Quand un cluster est détecté, les personnes travaillant dans l'entreprise, l'école ou l'université sont vaccinées (si elles le désirent) ainsi que les personnes vivant avec elles. Nous avons comparé cette stratégie avec des vaccinations non réactives (soit de masse, soit ciblant les lieux de travail, les écoles ou les universités). Nous avons montré que, dans la plupart des scénarios, la vaccination réactive était plus performante que les stratégies non réactives pour atténuer l'épidémie à nombre égal de doses vaccinales. Une approche combinée, avec une vaccination de masse et une vaccination réactive, était également plus efficace que la seule vaccination de masse. Nous avons testé l'impact des paramètres d'entrée sur l'efficacité de la vaccination réactive afin de comprendre dans quelle situation cette stratégie de vaccination était indiquée. Dans l'ensemble, les paramètres affectant le nombre de personnes vaccinées autour d'un cas ont un impact plus important sur l'efficacité de la vaccination réactive. Par exemple, si le nombre de personnes vaccinées est élevé dès le début de la simulation, il y aura moins de candidats à la vaccination réactive autour d'un cas détecté. Dans ce cas, la vaccination réactive sera moins avantageuse que lorsque moins de personnes sont initialement vaccinées. Inversement, si l'on modifie le nombre de reproducteurs, il y aura moins de candidats à la vaccination réactive autour d'un cas détecté. L'immunité naturelle au départ ou la réduction des contacts en raison du télétravail ou de la limitation des activités sociales n'ont eu que peu d'impact sur l'avantage relatif de la vaccination réactive. L'absorption du vaccin a eu un impact important sur l'efficacité de la vaccination réactive lorsqu'elle vise à la fois l'atténuation et le contrôle de la maladie.

Publications inclues dans cette thèse

[3] **Faucher, B.**, Assab, R., Roux, J., Levy-Bruhl, D., Tran Kiem, C., Cauchemez, S., Zanetti, L., Colizza, V., Boëlle, P-Y and Poletto C. "Agent-based modelling of reactive vaccination of workplaces and schools against COVID-19", *Nature Communications* **13**, 1414 (2022).

[1] **Faucher, B.**, Sabbatini, C. E., Czuppon, P., Kraemer, M. U. G., Lemey, P., Colizza, V., Blanquart, F., Boëlle, P-Y. and Poletto C. "Drivers and impact of the early silent invasion of SARS-CoV-2 Alpha", *Nature Communications* **15**, 2152 (2024).

Bibliography

1. Faucher, B., Sabbatini, C. E., Czuppon, P., *et al.* Drivers and impact of the early silent invasion of SARS-CoV-2 Alpha. *Nature Communications* **15,** 2152. https://www.nature.com/articles/s41467-024-46345-1 (2024).
2. Shu, Y. & McCauley, J. GISAID: Global initiative on sharing all influenza data – from vision to reality. *Eurosurveillance* **22,** 30494. https://www.eurosurveillance.org/content/10.2807/1560-7917.ES.2017.22.13.30494 (2017).
3. Faucher, B., Assab, R., Roux, J., *et al.* Agent-based modelling of reactive vaccination of workplaces and schools against COVID-19. *Nature Communications* **13,** 1414. https://www.nature.com/articles/s41467-022-29015-y (2022).
4. Béraud, G., Kazmercziak, S., Beutels, P., *et al.* The French Connection: The First Large Population-Based Contact Survey in France Relevant for the Spread of Infectious Diseases. *PLOS ONE* **10,** e0133203. https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0133203 (2015).
5. *google mobility data - Recherche Google* https://www.google.com/search?q=google+mobility+data&oq=google+mobility+data&aqs=chrome..69i57j0i22i3019.3715j0j4&sourceid=chrome&ie=UTF-8.

SORBONNE UNIVERSITÉ

DOCTORAL THESIS

Modelling COVID-19 spread to reconstruct the viral dissemination and inform the design of interventions.

Author:
Benjamin FAUCHER

Supervisors:
Pierre-Yves BOËLLE,
Chiara POLETTO

Laboratory:
Equipe SUMO, IPLESP UMR-S 1136, INSERM

A thesis submitted for the degree of Doctor of Philosophy in Epidemiology and Public Health

Ecole Doctorale Pierre Louis de Santé Publique:
Epidémiologie et Sciences de l'Information Biomédicale

Abstract

Emerging pathogens pose significant challenges to public health authorities. In the context of the COVID-19 pandemic, the SARS-COV-2 and the variants of concern followed a similar pattern. A new virus emerged in one country, spread globally, and then triggered a rapid surge in cases worldwide. To deal with this situation, it is critical to monitor the epidemic, decipher incomplete and incoherent data, and rapidly design interventions. Mathematical models can help interpret heterogeneous surveillance data and inform the design of interventions. In this thesis, we addressed both aspects. First, we developed a mathematical framework to understand how surveillance and epidemic drivers concur in shaping observations. We retrospectively reconstructed the international spread of the Alpha variant in the Fall of 2020 from sequencing and air travel data. In a second work, we focused on intervention. We proposed an agent-based model to quantify the epidemiological impact of a reactive vaccination strategy targeting workplaces and schools where cases are detected. We tested the effectiveness of this strategy to mitigate a general rise in cases and to limit the spread of a new variant.

Acknowledgment

I would like to thank the people who contributed to the success of the three years that led to the writing of this thesis. First, I want to thank the two reviewers of this thesis, as well as the members of the jury, for taking the time to read the manuscript and for providing me with both insightful and relevant comments and remarks.
I arrived in Paris in early 2021, in the midst of the COVID pandemic, to begin an internship in the laboratory where I subsequently completed my thesis. Looking back, I remain amazed at the luck I had in meeting exceptional people throughout these three years. It is these people that I would like to sincerely thank. My first thoughts go to my two supervisors, who each contributed greatly to my work. First to Chiara, who, through her scientific and human qualities, allowed me to progress in this thesis and grow professionally. Above all, her precision in both data analysis and writing, as well as her ability to find the delicate balance between supervision and autonomy, makes her an example for me in my future career. Then to Pierre-Yves, my thesis director, who always has the right advice to unblock any situation, regardless of the difficulty.
I would also like to thank the rest of the SUMO team at IPLESP. First, Vittoria Colizza, Raphaëlle Metras, and Eugenio Valdano, for the spirit they instill in this fantastic team where a good atmosphere and high standards mix. I would then like to thank my colleagues, whether or not they are still in the laboratory: Laura, Giulia(s), Elisabetta, Chiara, Mattia, Wen, Francesco, Marika. I have a special thought for the "canteen club": Davide, Anthony, Pourya, and Francesco.
I would also like to thank the people I met outside the laboratory, who allowed me to have a balanced life and also helped me develop over these past three years. First, Aymeric, Adrien, and FX, who, in addition to becoming very close friends, are examples to me through their human qualities, including honesty, kindness, and intelligence. But also Brice, Massy, Eva, Léna, Antoine, Basile, Victor, and Hélène.
Finally, I would like to thank my parents and brothers for their daily support.

Contents

List of Figures

Preface

The Covid-19 pandemic has been characterised by successive waves of cases. Whether it is the original strain that emerged in China or the subsequent variants, the pattern has repeated itself: a new virus or variant emerged in one country, spreads globally, and then triggered a rapid surge in cases worldwide. At each stage of the dissemination, it is crucial to monitor the epidemic unfolding to get the necessary information to plan tailored interventions. Mathematical models are essential tools in this effort. Appropriate models help interpret the data coming from heterogeneous surveillance systems. They allow for reconstructing the virus spread by inferring unseen variables. However, understanding the epidemic situation is not enough. Models can be used to simulate different interventions in silico. This allows for the testing of various scenarios and quantifying the cost-benefit balance of these interventions. Policymakers can then use the results to implement public health measures.

Since the onset of the COVID-19 pandemic, many cases have gone undetected by surveillance systems for several reasons. Clinical symptoms were aspecific and could easily be mistaken for those of other respiratory illnesses such as influenza. Many cases remained entirely asymptomatic. In addition, the progression of the disease featured a presymptomatic phase during which individuals were infectious but displayed no symptoms, delaying the detection. Ensuring massive virological surveillance was challenging for practical reasons. Early in the pandemic, case detection only relied on clinical symptoms until the first PCR tests became available in mid-January 2020 [1]. While the capacity to detect COVID-19 cases increased throughout 2020, the emergence of new variants posed new challenges as their detection and monitoring required the implementation of sequencing systems. These systems demanded resources and infrastructures that were less available in low-income countries, which caused strong heterogeneities in global virological surveillance [2].

Initially, travel bans aimed to contain the virus at its source and prevent importation. In January 2020, all travel to and from Wuhan was prohibited [3] and a similar approach was used with the emergence of new variants. Many countries restricted travel from the UK, South Africa, and India, where Alpha, Beta, and Delta variants emerged [4]. Measures such as active surveillance or contact tracing attempted to prevent imported cases of SARS-COV-2 and successive VOCs. Once it was possible, reactive vaccination was also used to prevent the spread of the Delta variant. However, due to the underdetection of initial cases, the virus or the new variant silently spread and sparked local outbreaks in other countries. Consequently, authorities were forced to implement mitigation measures designed to curb local transmission - lockdowns, curfews, social distancing, contact tracing and, when available, vaccination.

In this thesis, I carried out modelling work dealing with both sides of the medal: monitoring and intervention. It is first necessary to understand the epidemic situation to properly respond to the spread of a new virus or variant. This requires the ability to interpret surveillance data correctly. In the first part of this thesis, we focused on this aspect. We retrospectively analysed the spread of the Alpha variant out of the UK [5]. We sought to relate the observed invasion to the actual one and quantify the detection delays caused by the limited and heterogeneous surveillance. In the second part of

this thesis, we focused on assessing an intervention measure to contain and mitigate a new variant once it is detected [6]. In particular, we studied the reactive vaccination. We have developed an agent-based model to quantify the potential of this strategy to mitigate the epidemic and limit introductions to spread locally further. We did this study as the vaccination campaign unfolded in the first semester of 2021 to support public health authorities' recommendations [7].

In Chapter 1 , I provide an overview of the epidemiological context: how SARS-CoV-2 emerged in China in 2019 and subsequently spread worldwide. I discuss new variants' emergence, dissemination, and impact on the pandemic dynamics. Then, I examine how authorities have implemented surveillance systems to detect new outbreaks or the arrival of a new variant. I further explore the measures implemented to contain the spread of a new virus or variant or mitigate its impact once established.

In Chapter 2, I review models used during the COVID-19 pandemic. I first focus on the dynamical models used to study the early international spread of SARS-COV-2 and the subsequent emergence and spread of new variants. Secondly, I provide an overview of different modelling frameworks used to quantify the effectiveness of interventions. I present modelling frameworks developed before the pandemic and how they were used for SARS-COV-2.

In Chapter 3, I present the retrospective study on the spread of the Alpha variant during its emergence in the second half of 2020. I first introduce the data, including metadata from sequences submitted to GISAID and air travel data. By comparing them with other studies, I show why these surveillance data suggested the existence of silent spread. Then, I present the results published in [5]. In this study, we reconstructed the silent spread of the Alpha variant and quantified its duration. We then investigated the factors that contributed to this silent circulation. Finally, we simulated the local spread of the Alpha variant in six countries with two different models [8] [9, 10, 11]. We recovered the Alpha frequency obtained in the national investigation. These results provided a validation of our model.

In Chapter 4, I present the results of our study on reactive vaccination, published in [6]. First, I discuss the opportunities and challenges of reactive vaccination, drawing on historical examples. Then, I present the agent-based model used in our analysis. I detail the framework that produced a synthetic population from demographic and census data and how we used it to study reactive vaccination. We conducted this study in the context of a rise in cases caused by the importation of the Delta variant in France. We quantified the ability of reactive vaccination to limit the initial spread of the variant. We also explored the ability of this strategy to mitigate a widespread epidemic. We tested different scenarios and examined the influence of various factors (initial incidence, population immunisation, social distancing measures). We compared this strategy to mass and other targeted vaccination strategies.

In the conclusion, I draw conclusions to summarise the results presented in this thesis and discuss potential avenues for further analysis to complement this work.

Publications included in this thesis

[6] **Faucher, B.**, Assab, R., Roux, J., Levy-Bruhl, D., Tran Kiem, C., Cauchemez, S., Zanetti, L., Colizza, V., Boëlle, P-Y and Poletto C. "Agent-based modelling of reactive vaccination of workplaces and schools against COVID-19", *Nature Communications* **13**, 1414 (2022).

[5] **Faucher, B.**, Sabbatini, C. E., Czuppon, P., Kraemer, M. U. G., Lemey, P., Colizza, V., Blanquart, F., Boëlle, P-Y. and Poletto C. "Drivers and impact of the early silent invasion of SARS-CoV-2 Alpha", *Nature Communications* **15**, 2152 (2024).

Chapter 1

Covid-19 context: epidemic dynamics, surveillance and intervention.

In this chapter, I bring an overview of the epidemiological context. I discuss the general dynamic of the spread of Sars-Cov-2 and the different variants. I also present how countries have implemented monitoring systems and undertaken efforts to curb the spread of the virus.

1.1 Chronology of COVID-19 invasion.

1.1.1 Coronavirus disease 2019.

COVID-19 was an infectious disease caused by the Sars-Cov-2 virus, mainly transmitted through respiratory droplets during close contact between individuals and via aerosols [12, 13]. It was characterised by symptoms common to other respiratory illnesses [13], including fever, fatigue, cough, sore throat, and difficulty breathing. The illness varied in severity, from completely asymptomatic cases to severe pneumonia necessitating hospitalisation, sometimes in intensive care units, and even leading to death. Analyses of early cases in Wuhan revealed that 13.8% were severe, with 6.1% classified as critical [13]. The probability of developing severe symptoms increased with age and the presence of comorbidities (hypertension, diabetes, cardiovascular disease, chronic respiratory disease, or cancer). An analysis of reported cases in France [14] estimated that 2.9% of infected cases required hospitalisation, with a 0.5% fatality rate. This fatality rate varied nearly 1,000-fold by age, ranging from 0.001% for individuals under 20 to 8.3% for those over 80.

The pandemic initially emerged in China in Wuhan and spread globally in early 2020. The first recorded patient was hospitalised with pneumonia on December 12, 2019 [15]. By January 30, 2020, the World Health Organization (WHO) confirmed 7,818 cases worldwide, with 82 cases in 18 countries outside China [16]. On March 11, 2020, WHO declared COVID-19 as a pandemic [16]. The first case was reported in France on January 26, 2020 [17, 18]. The pandemic's progression has been marked here by successive waves of infections [19]. The first wave occurred in the spring of 2020, curved by the lockdown, followed by low virus circulation in the summer. A resurgence in early fall 2020 marked the onset of a second wave, peaking in early November and sustaining transmission throughout winter. This last dynamic resulted from a series of restrictions during the autumn (closure of bars and public places, followed by curfews and lockdown), combined with the emergence of the Alpha variant, which was more transmissible. A third wave occurred during spring 2021 with the emergence of the Delta variant. In the following months, as the vaccination campaign was ramping up,

social distancing measures were gradually relaxed. In January 2022, the emergence of the Omicron variant triggered a new wave of cases.

1.1.2 Propagation of Sars-Cov-2 across the globe.

After emerging in Wuhan, SARS-CoV-2 spread across China and other continents [18, 20, 21]. Different studies suggested that the spread of SARS-CoV-2 was largely mediated by international travel. At the within-country scale, the analysis of the spatial distribution of cases detected around Wuhan showed that the main factor in the spread of the epidemic was the flow of population from Wuhan [20]. The reconstruction of the virus dissemination based on international travel data gave consistent results with phylogenetic analyses [22, 23, 24]. In January, researchers used air traffic data to Europe to estimate the risk of Covid-19 importation [25]. Their findings were consistent with the epidemic situation observed a few weeks later. In particular, based on airline traffic data, they estimated that the UK, Germany, France, Italy and Spain faced the highest risk of importation in Europe. Indeed, these countries were the first to report clusters [26]. Combining phylogenetic analysis, international traffic data and analysis of imported cases allowed researchers to reconstruct the initial propagation of SARS-COV-2 from Wuhan. The international dissemination was found to occur in two phases [21, 22]. During the first phase, cases were exported from Wuhan to cities across Europe, Asia and North America [21, 23]. Containment measures and travel bans implemented in China in late January drastically declined exportation [17, 18]. However, most of the cases exported during this first phase went undetected. If 288 cases and 42 clusters were identified outside of China between January 3rd 2020 and February 13th 2020, between 36% and 65% of imported cases until mid-February may have gone undetected [18]. Then, cases continued to spread from countries outside of China silently. Europe became the main epicentre of the epidemic in February 2020 [22], fueling a second phase of the global dispersion of the virus to the USA and other countries [23].

1.1.3 Variants of Concern.

The emergence of new variants contributed to shaping the dynamic of the pandemic. Since SARS-CoV-2 is an RNA virus [27], it undergoes frequent mutations. Many of these mutations were found in the Spike glycoprotein. Some of them resulted in phenotypic changes, altering the virological and epidemiological characteristics of the virus (new strains). Some of these mutations provided an increased transmissibility or an immune escape. Various classification systems were used to identify new variants. The Pango Lineage system [28] classified variants based on lineages and sub-lineages. This classification relied on the variant's phylogenetic tree. Lineages were named using a prefix with a letter (B or BA) followed by a number (.1 or .1.1.5) with additional numbers added to the nomenclature for sub-lineages. Several surveillance organisations, including the CDC [29], ECDC [30], and WHO [31], employed a classification system based on the level of concern associated with the lineage: Variant of High Consequences (VOHC), Variant of Concern (VOC), Variant of Interest (VOI), and Variants Being Monitored (VBM) [19, 29]. A VOHC had clear evidence that prevention measures or medical countermeasures had significantly reduced effectiveness relative to previously circulating variants. This includes failure of diagnostic tests, a significant reduction in vaccine effectiveness, a reduced susceptibility to therapeutics or a higher hospitalisation rate. A VOI was a variant that exhibited a mutation known or predicted to change its characteristics and was likely to impact public health (increased severity,

reduced treatment effectiveness). When the correlation between the new phenotype and the epidemiological impact had been demonstrated, the VOI was classified as a VOC. VBM was the lowest level of severity of monitored variants. This includes variants with possible impact on public health based on genomic data or former VOI/VOC not circulating anymore.

As of May 2021, WHO has used a nomenclature based on Greek letters to facilitate communication about the most monitored variants (Alpha, Beta, Gamma, Delta, Omicron) [32].

1.1.4 Dissemination of Variants of Concern.

Several VOCs have emerged throughout the pandemic, appearing in one country and spreading globally. Some have replaced previous variants, while others coexisted [21]. The first variant, retrospectively identified, emerged in January 2020, with a mutation at the D614G site [21, 33] and was associated with increased transmissibility during the surge of cases in Europe in early 2020 [33].

The Alpha variant (B.1.1.7) was initially detected in the United Kingdom in late September 2020 [30, 34], exhibiting increased transmissibility compared to the historical strain [21, 35, 36]. The Beta variant (B.1.351) was first detected in early October 2020 in South Africa [30, 34], and it showed both increased transmissibility and immune evasion [21, 37]. The Gamma variant (P.1) was first detected in Brazil in December 2020 [30] and exhibited immune escape [21]. The Delta variant (B.1.617.2), initially detected in India in March 2021 was more transmissible and showed immune escape [30]. The Omicron variant (B.1.1.529), first identified in South Africa in late 2021 [30], exhibited higher transmission and immune evasion [30]. It gradually replaced earlier variants. Phylogenetic modelling of sequenced viruses worldwide has enabled the reconstruction of the historical spread of different variants across the globe [38, 39, 40, 41]. The initial emergence of the first three VOCs (Alpha, Beta, and Gamma) led to their simultaneous circulation, replacing the historical strain. In this context, the countries of origin (the UK for the Alpha variant, South Africa for the Beta variant, and Brazil for the Gamma variant) served as the primary source of case importation into other countries during the initial phase. The Alpha variant initially disseminated from the UK to Europe, to Asia and North America, where it subsequently fueled the epidemic. Similarly, the Beta variant first spread from South Africa to North America, and then to Europe, and Asia by the end of 2020. Asia became, after that, a secondary source of this variant in April 2021. Meanwhile, the Gamma variant remained largely confined to South America until its exportation to other continents around May 2021. In the second quarter of 2021, the Delta variant replaced these three VOCs [38]. After causing a large outbreak in India in March 2021 [41], the Delta variant spread globally. Unlike previous variants, India was not the major source of importation of the Delta variant. Less than 15% of all introductions to other countries were attributed to India, other countries rapidly acting as secondary sources [38]. The same pattern was observed for the two lineages of the Omicron variant (BA.1 and BA.2) after their emergence in South Africa. After one week of global dissemination, countries outside of South Africa were already the source of importation [38]. Data from sequencing obtained through the Flash Surveys conducted by Santé Publique France [19] have revealed four waves of variants in France (Figure 1.1), playing a critical role in the dynamic of the epidemic. The historical strain dominated throughout 2020. The first cases of Alpha variants were detected in the last days of December 2020. In the following months, an increase in cases occurred despite the enforcement of stringent measures. Curfews and social distancing helped to limit the spread of the historical strain but were not strong enough

to contain the exponential rise in cases caused by the new variant [10]. This propagation led to a plateau of cases, followed by a new wave. In April 2021 the Alpha variant accounted for over 80% of sequenced cases. Then, it was progressively replaced by the Delta variant in June 2021 (see Figure 1.1). From the December 6, 2021 survey, the Omicron variant gradually replaced the Delta variant and represented 100% of sequenced cases in surveys from February 14, 2022.

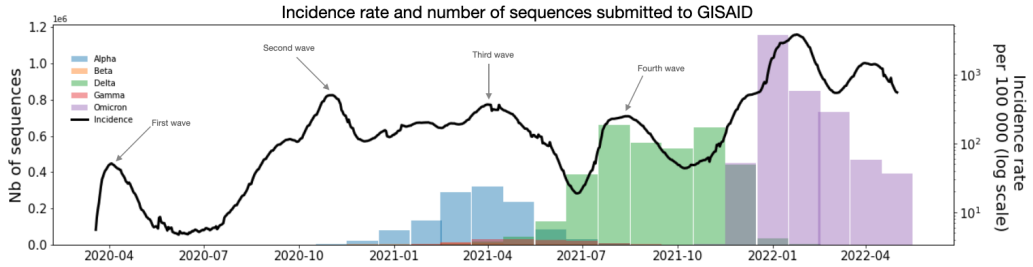


FIGURE 1.1: Evolution of the pandemic. The black line (right y-axis) represents the recorded incidence in France. Values are extracted from French government dashbord [42]. Coloured bars represent the number of sequences submitted to GISAID by Variant of Concern [43].

1.2 Response to the epidemics of Sars-Cov-2 and VOCs.

1.2.1 Monitoring of cases and detection of new outbreaks.

Following the emergence of SARS-CoV-2 in China in late 2019, other countries worldwide implemented strategies to detect imported cases. In Europe, the ECDC and the WHO regional division recommended that European countries report all detected cases based on WHO criteria [26]: a confirmed case by RT-PCR regardless of symptoms or a suspected case with a negative laboratory test. The investigation of early cases revealed the existence of clusters in France and Germany. Subsequently, various surveillance systems were established with different objectives [44, 45]: following the virus spread, conducting virological surveillance, identifying at-risk groups, monitoring the impact on healthcare systems, and assessing the effects of mitigation measures. Genomic surveillance of SARS-CoV-2 has resulted in unprecedented sequencing coverage. Since the onset of the pandemic, over 16 million SARS-CoV-2 sequences have been collected and submitted to the GISAID platform [43]. However, sequencing coverages varied over time and between countries. Countries with lower sequencing capacity may have identified a new variant with a delay of several weeks or months [46]. French authorities, notably Santé Publique France (SPF), rapidly established surveillance of variants circulation in France, aiming at achieving two primary objectives. The first objective was to continuously measure the relative prevalence of different variants within the French population in real-time. The aim was to understand the epidemic situation and model variants' impact on the overall epidemic dynamics. The second objective was the early identification of variants of interest.

The first step of variant detection was screening [19]. When a COVID-19 case was confirmed via RT-PCR, screening could rapidly identify specific known mutations within 24 hours. Full genome sequencing of the viral genome allowed for a more complex genetic analysis, capable of detecting any variant, even unknown ones. Unlike screening, however, this method yielded results within a timeframe ranging from one to two weeks. This sequencing served three distinct purposes:

- *Representative Sequencing*: The goal was to obtain a representative sample of variant diversity within a target population through regular Flash Surveys. These flash surveys were coordinated by the EMERGEN consortium [47, 48], gathering multidisciplinary experts to roll out a nationwide genomic infection surveillance system. These surveys involved a network of public and private laboratories capable of using Next Generation Sequencing (NGS) techniques to fully sequence detected SARS-Cov-2 cases. Ces enquêtes Flash ont été mises en place pour cartographier à un instant donné les variants du SARS-CoV-2 circulant en France. In the first two surveys (January 7-8, 2021, and January 27, 2021), only samples previously screened as Variants of Concern were sequenced. Subsequent surveys included sequencing from a random sample of SARS-CoV-2 positive samples, regardless of the screening result. To maintain sample representativeness, the samples were not drawn from cluster investigations.

- *Targeted Sequencing*: The objective was to perform more comprehensive and systematic sequencing within a specific target population, such as severe hospital cases, immunocompromised patients, or treatment failures.

- *Interventional Sequencing*: This reactive sequencing addressed a specific local situation, such as a rapid increase in cases or hospitalisations, to detect potential clusters of an emerging variant.

1.2.2 Containment measures.

Just after the emergence of SARS-CoV-2 or a new variant, the primary goal of health authorities was to contain the epidemic at its source and prevent the importation of cases. On January 23, 2020, Chinese authorities banned all travel from and to Wuhan to prevent the spread of the novel Sars-Cov-2 to other provinces of China [3]. In parallel, many other countries imposed travel reductions from China to prevent the importation of cases. Many airline companies limited or suspended their flights to China [17]. With the emergence of new variants, a travel ban from the source country was also used to prevent importation in other countries. In early December, many European countries banned travel from the UK after the Alpha variant emergence [4]. The same restrictions were applied to South Africa for the Beta and Omicron variants [49] and to India for the Delta variant [50]. Strict travel bans were combined with other restrictions on travellers such as quarantine and screening [51]. The effectiveness of these measures was limited. Studies of past epidemics have already shown that reducing air traffic has a limited impact on the spread of a virus. In the case of H1N1 influenza, a 40% drop in travel flow from and to Mexico led to a delay of less than three days on average in the arrival of the virus in other countries [52]. With a reduction of 90% in travel the delay would have been around two weeks. In 2014, airline traffic connecting the West Africa region affected by Ebola was reduced by 60%. This was not sufficient to prevent the exportation of the virus and resulted in a delay of only a few days [53]. Containment measures in China for Sars-Cov-2 in January 2020 were much more intense and drastically decreased the reexportation of cases [17, 18, 54]. However, these measures were implemented too late, and countries outside China witnessed an increasing report of clusters in early February [18]. The problem was similar with the following variants, which circulated silently before the implementation of measures [38]. Once the variant circulates locally, travel measures were found to be ineffective [51, 55].

Contact tracing strategies consisted of detecting cases and tracing their contacts to break transmission chains. Detected cases and their contacts were then tested and possibly quarantined. The aim was to target infected individuals to avoid costly general interventions such as lockdowns. This strategy worked well in controlling the

epidemic in some places like Singapore, South Korea and China but failed in other countries [56]. Contact tracing policy was shown to need high tracing coverage, strong quarantine and short contact tracing delays. The contact tracing policy was implemented in France after the first wave to prevent the reintroduction of cases during the summer of 2020 [57]. However, this was not strong enough, with an average of 2.1 cases detected per index case [57], to prevent a rise in cases in September and the implementation of a second lockdown.

1.2.3 Mitigation measures.

With containment efforts failing, most countries have implemented mitigation measures to prevent a collapse of the healthcare system. If travel restrictions and contact tracing failed to contain the spread of the virus, they participated in mitigating the epidemic [18, 58, 59]. Mitigation resulted from a combination of implemented measures and responses to the rise of cases which aimed to limit the local transmission. The Oxford COVID-19 Government Response Tracker project (OxCGRT) documented the measures taken by different countries worldwide [60]. At some point during the pandemic, over 80% of countries adopted measures such as the closure of public places (schools, workplaces, public transports), cancellation of public events, restrictions on gatherings, restrictions on internal movement, stay-at-home requirements and mask mandates.

In France, schools were closed on March 12, and the first national lockdown was implemented on March 17, 2020 [61]. These measures were gradually lifted starting on May 11. On October 14, in response to a resurgence of cases, a curfew was imposed, followed by a second lockdown. From December 15, 2020, the lockdown was lifted, but the curfew remained. A third lockdown was implemented for four weeks on March 18, 2021, to curb the spread of Alpha. Initially at the regional level and then nationalised, this lockdown was less stringent than the previous ones. During the first lockdown, schools were closed as well as all non-essential businesses. People were required to stay at home, except for essential needs (groceries or medical) or a daily leisure time of one hour within a one-kilometre radius of their residence [61]. During the second lockdown, schools remained open, as well as more workplaces [62]. During the third lockdown, stay-at-home requirements were less stringent: people could move freely in a radius of ten kilometres around their homes [62]. Mask-wearing became mandatory after the first wave, initially in all public places [63] and later only indoors. Facilities hosting public (museums, theatres, bars, restaurants, gyms) and outdoor gatherings were alternatively closed [63] or subjected to capacity limits. Teleworking was made mandatory and then only encouraged to minimise contact in workplaces and in public transport [64]. The CoviPrev surveys [65], conducted by Santé Publique France (SPF), reported that a majority of the French population adopted social distancing measures. These measures included ventilating rooms, greeting without shaking hands, avoiding hugs, regular handwashing, wearing masks in public, and avoiding gatherings and face-to-face meetings. The survey results throughout the pandemic are presented in Figure 1.2.

Many studies were conducted to assess the effectiveness of mitigation interventions. Effectiveness on outcomes such as the number of severe cases or deaths has been difficult to estimate. The delay at the individual level between the infection and the onset of critical symptoms or the death results in a lag between the implementation of measures and the observed effect on these types of outcomes. Studies on the effectiveness of the transmission rate are more reliable [66]. The reduction of transmission rate

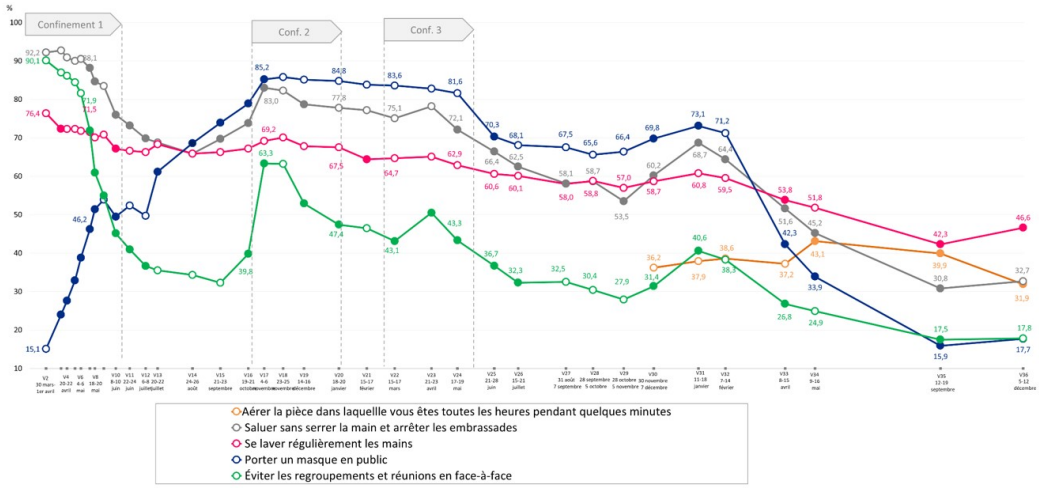


FIGURE 1.2: Adoption of social distancing measures since the start of the pandemic, as measured by the Coviprev surveys. Figure from [65].

was shown to mostly results from a combination of measures, explaining the high effectiveness of lockdown [67]. Studies estimating the impact of each measure suggested that limiting gathering and school closure were the most effective measures, while additional stay-at-home orders and border closures had a limited impact on transmission rate [68, 69].

1.2.4 Vaccination.

The introduction of several COVID-19 vaccines in late 2020 was a game-changer in the fight against the pandemic. Vaccines relying on mRNA technology were rapidly available and clinical trials demonstrated good efficacy. In particular, they showed high effectiveness in preventing symptomatic and severe forms of the disease. Their deployment was then a critical tool to alleviate the burden on healthcare systems and to finally ease social distancing measures. They also helped to prevent infections and limit the spread of new variants or outbreaks.

The initial studies reported an efficacy of 95% against symptomatic COVID-19 seven days after the second dose of the BNT162b2 mRNA vaccine provided by Pfizer-BioNTech [70]. Studies regarding the mRNA-1273 vaccine provided by Moderna reported a similar efficacy of 94.1% 14 days after the second dose for the historical strain. The ChAdOx1 nCoV-19 vaccine provided by AstraZeneca showed lower efficacy, with effectiveness against symptomatic cases of 81.7% for the historical strain and 70.4% for the Alpha variant.

These estimates were computed from pre-vaccination campaign studies. After a few months, they needed refinement for new variants, booster doses, and waning immunity. A meta-analysis published in 2023 [71] provided an overview of the effectiveness of mRNA vaccines during the pandemic. For mRNA vaccines, protection against symptomatic infection was shown to decrease over time. Effectiveness, across all variants, decreased from 87% [84%, 90%] 14 days after a full vaccination schedule (one or two doses, depending on the vaccine) to 66% [57%, 74%] after 112 days and dropped to 51% [22%, 69%] after 280 days. However, effectiveness against hospitalisation and mortality remained high over time. For hospitalisation, effectiveness decreased from 93% [89%, 95%] after 14 days to 80% [64%, 88%] after 224 days. The

mortality, decreased from 94% [88%, 97%] after 14 days to 87% [73%, 94%] after 140 days.

The emergence of new variants also affected vaccine effectiveness [21], as initial vaccines were designed for the historical strain. Protection against symptomatic infection after 14 days remains high for the Delta variant (91% [88%, 93%]) but is reduced for the Omicron variant (67% [53%, 77%]). Updated vaccines have recently shown high effectiveness against Omicron and its subvariants [72].

Despite the rapid deployment of COVID-19 vaccines worldwide, with the first doses distributed in early December [73], vaccine coverage in most Western countries was far below the targets set by authorities at the start of summer 2021. As of June 1, 2021, six months after the vaccination campaign's launch, the United States had distributed 315.29 million doses, France 38.23 million, and Canada 24.38 million. Assuming that each vaccinated individual receives two doses, this represents a vaccinated population share of 47% for the USA, 28% for France, and 31% for Canada.

When implementing vaccination policies, public health authorities had to face practical challenges that limited or slowed down the effectiveness of vaccine deployment. In France, supply issues, logistical implementation, and public acceptance hindered the rapid establishment of optimal vaccine coverage against COVID-19. While the Pfizer vaccines allowed the vaccination campaign to begin in December 2020 [74], their distribution required the purchase and installation of ultra-cold freezers. Mass vaccination required the receipt of vaccine doses, organisation of various delivery methods, transportation of doses, as well as all accompanying equipment (needles, syringes, sterile medical devices) to pharmacies and equitable distribution throughout the country.

To control the epidemic more efficiently despite limited doses, most of the countries opted to target populations at the highest risk of severe COVID-19 outcomes. On December 27, the first vaccination phase [75] was launched, following the recommendations of the French Haute Autorité de Santé (HAS) and supported by modelisation [76]. This initial phase targeted elderly care facilities and healthcare personnel. Then, vaccination was available progressively to other parts of the population. The main vaccination points included elderly care facilities, healthcare facilities, pharmacies, and vaccination centres and used doses from Pfizer, Moderna and AstraZeneca.

1.3 Conclusion.

In this chapter, I reviewed the spread of SARS-COV-2, the surveillance efforts, and the intervention response to the virus. The underdetection of cases led to a poor understanding of the virus spread in the early phase of the pandemic. The decrease in cases observed in China in January 2020 following the implementation of lockdown measures created a false sense of security, making it difficult for health authorities of other countries to justify the implementation of restrictive measures. With the virus already circulating locally, travel bans and border closures had little impact on the local dynamic of the epidemic. The same problem occurred with the emergence of VOCs. For instance, in the case of the Alpha variant, European countries banned travel in December 2020 [4] but the Alpha variant was likely circulating locally at that time and caused a rise in cases during the first months of 2021. The COVID-19 pandemic provided an unprecedented amount of data on the virus's spread and detection. We need to exploit it and learn from COVID-19 experience to minimise the impact of underdetection in future pandemics. Retrospective studies can help to understand the drivers

of cryptic transmission and its epidemiological consequences. To do so, complex models are needed to integrate the heterogeneous data available from surveillance systems. Understanding the epidemic situation is an essential step to plan interventions, but is not enough and policymakers must rapidly implement effective measures. However, all measures come with costs, some very high like lockdowns. It is, therefore, crucial to evaluate the cost-benefit balance of these interventions. This work involves estimating the impact of a measure before implementing it. As conducting real experiments is not feasible, mathematical models enable the testing of various scenarios to compare the effectiveness of various interventions or combinations of interventions. However, uncertainties remain about the epidemiological situation which translates into uncertain and hard-to-estimate parameters. It is essential to quantify how different parameters affect the effectiveness of these interventions. In the following chapter, I will present the models used during the pandemic to study the international spread of the virus, as well as those used to assess the effectiveness of intervention strategies.

Chapter 2

Modelling frameworks to study the COVID-19 epidemic.

The COVID-19 pandemic has been accompanied by intense scientific production [77], in which modelling played a significant role. Modelisation has been used for epidemic assessment, estimating key epidemiological parameters (reproduction number, generation time, part of asymptomatic...), nowcasting and forecasting the epidemic unfolding, and assessing the effectiveness of intervention strategies. Depending on the specific objective, modelling studies can be prospective or retrospective, theoretical or data-driven. During a pandemic, modelling outputs must be concise and actionable, tailored to the specific needs of policymakers. Furthermore, models must be adapted to the available data for parametrisation and validation [78]. Modelling efforts during the COVID-19 pandemic used frameworks developed for past epidemics and adapted to the current situation. We present here a review of modelling methodologies employed during the COVID-19 pandemic to study early dissemination and assess the impact of interventions.

2.1 Modelling global invasion.

Several methods have been employed to study the global spread of new viruses or variants, with several objectives. At the onset of a pandemic, the aim was to understand the epidemiological situation, identify the epicentre of the epidemic, and determine the countries that had already imported the virus. In the short term, propagation models allowed for estimating the risk of virus exportation. Studying the first imported cases also enabled the quantification of relevant characteristics to understand the epidemic: detection rate, effective reproduction rate, proportion of asymptomatic cases, or generation interval. Analysis can also be retrospective to reconstruct the pattern of global invasion. Some studies relied on the assumption that the risk of importation into a country was proportional to the volume of air traffic between the outbreak source and the country. Before the COVID-19 pandemic, air transport data had already been used to estimate the risks of exportation of MERS [79, 80] and Ebola [53] among others. During the COVID-19 pandemic, the same approach was employed at the onset of the pandemic to estimate the risk of the virus importation from the top ten cities in China receiving the most air passengers [81]. More detailed analyses focused on the risk of importation in African [82] and European countries [25]. Air traffic data were also used to analyse the spread of variants. In [83], the authors compared the spatial distribution of the Alpha, Beta, and Gamma variants at the end of February 2021 with the distribution of major traveller destinations from the source countries of these variants during the fall of 2020 (namely the UK, South Africa, and Brazil). Mobility data from the UK was used to estimate the risk of exportation of the Alpha variant in fall

2020 [84]. Additionally, air traffic data from the UK and South Africa were combined with phylogenetic analyses to trace the spread of the Alpha and Beta variants in October 2020 [34]. At the onset of the pandemic, recorded cases were sporadic, and only those reported by health agencies, government communications and media coverage could be used. Studying travelling cases allowed modelling of early growth and inference of the epidemic characteristic (detection rate, reproductive number, generation time...). In [85] authors used the number of detected cases and travel flows to estimate early in the pandemic the number of cases in Wuhan. In [18] authors collected early cases across multiple countries outside of China. They used a bayesian framework to predict the trends in importation and quantify the proportion of undetected imported cases. In [86] authors performed a regression between detected cases in locations outside of China and air travel volumes from Wuhan to estimate detection probability. In [87], the authors developed a framework based on the first reported cases to estimate the date of the emergence of Sars-Cov-2 or a new variant.

2.1.1 Metapopulation models.

Metapopulation models have been widely used to investigate the spread of a new virus or variant. These models describe the virus's propagation across interconnected subpopulations which can range in scale from city-level to country-level or continent-level. The type of connections (air traffic, road traffic, etc.) depends on the specific problem and model being used. The smallest metapopulation model consists of two interconnected patches, where a SIR model describes the transmission in each subpopulation. This model was used to theoretically study the interaction between two coupled areas [88, 89]. This framework has been expanded to describe larger networks as for the spread of the Hong Kong flu pandemic in 1968-1969 [90, 91, 92]. In these studies, the networks comprised major cities worldwide. The epidemic in each city was described by SIR or SEIR compartmental models. At each time step, a person in city i had a certain probability of travelling to city j. This probability was computed from the number of air passengers travelling between the two cities. With advancements in computational capabilities and the availability of big data, metapopulation models evolved to become more complex and realistic. Many large-scale metapopulation models were designed in the past, integrating real demographic and mobility data to simulate the global spread of epidemics [93, 94, 95, 96]. Incorporating stochastic temporal evolution also allowed for the description of intrinsic stochasticity in propagation and uncertainties in certain variables [89]. In [90, 97], the metapopulation model described a network of urban areas centred around 3100 airports connected by 17,182 links. In addition to being more comprehensive (representing 99% of global air traffic), this network better accounted for heterogeneities in transit capacity and traffic between countries. This framework served as the basis for the GLEAM (Global Epidemic And Mobility) project, a collaboration of researchers who leveraged mobility data integrated into stochastic models to simulate epidemic spread worldwide [98, 99, 100]. This project aimed at providing a realistic modelling tool for epidemic assessment in case of pandemics. It was used for various epidemics: Zika [101], Influenza [99], Ebola [53], seasonal influenza [102] and SARS-COV-2 [17, 103].

Besides providing the basis of computationally intensive simulations, the metapopulation framework allowed for analytical calculations to gain an understanding of the epidemic invasion dynamic. In [104] for instance, the authors computed the probability distribution of a virus's first arrival time for two connected subpopulations.

Let p be the probability to travel from city 0 to city 1 during a time interval Δt and $I(t)$ the number of individuals infected in city 0 at time t. The probability $P(t_1 = n\Delta t)$

that the virus arrives for the first time in $t_1 = nt$ is given by the probability of having no exportation of cases during $(n-1)t$ and at least one exportation at time nt :

$$P(t_1 = nt) = (1 - (1-p)^{I(t_1)}) \prod_{k=1}^{n-1} (1-p)^{I(k.\Delta t)}$$

Assuming that the growth of cases in city 0 is exponential ($I(t) = I_0 e^{at}$) and considering standard continuous approximation, the previous formula gives that t_1 follows a Gumbel distribution

$$P(t_1 = t) = p e^{at} . e^{-\frac{p}{a} e^{at}}$$

with a mean

$$< t_{\text{first introduction}} > = -\frac{1}{a} ln(\frac{p}{a})$$

This model was used to understand the impact of travel restrictions [55, 105]. In [52] authors analysed the effect of the 40% reduction in travel flows from and to Mexico during the H1N1 pandemic in 2019. They showed that this reduction led to an average delay of three days in the exportation of the virus [105]. Similar analyses have been conducted for other outbreaks, including SARS, the West Africa Ebola outbreak, and influenza epidemics [52, 53, 105, 106].

2.2 Modelling intervention.

Mathematical models are also powerful tools to test the effectiveness of interventions. Different types of models can be employed depending on the specific objective, timeframe, and available data. Most of these models were based on compartmental frameworks, like SIR or SEIR [107], where the population is divided into compartments describing the different states of illness (Susceptible, Exposed, Infectious and Recovered). Basic SIR or SEIR models assumed homogeneous mixing and described the evolution of the epidemic with a set of deterministic differential equations. However, deterministic models underestimate epidemic growth, particularly in the initial phase. Stochastic models were shown to be more effective in modelling outbreaks, as they account for the possibility of local epidemic extinction [8, 108]. It is also possible to relax the assumption of homogeneous mixing. Incorporating contact matrices based on demographic data allowed for stratifying the population into age groups [109, 110] to account for age-specific transmission patterns and the effects of social distancing across groups and settings.

At the onset of the COVID-19 pandemic, compartmental SIR or SEIR models, informed with disease-specific epidemiological parameters, were used to assess the impact of early interventions in China [111, 112, 113]. In [111], the authors analysed the reduction of contact in Wuhan and Shangai in early February 2020 compared with pre-pandemic surveys. From this data, they built an age-stratified SEIR model to assess the effectiveness of social distancing and school closure. They found that social distancing measures alone could stop the transmission, while school closure would not have been sufficient to control the spread. In [112], travel data combined with an SEIR framework was used to model outbreaks across China with social distancing. This study estimated that early detection and isolation of cases were more effective than social distancing measures in preventing infection. As more information on the infection became available, SIR or SEIR models have been refined by adding compartments

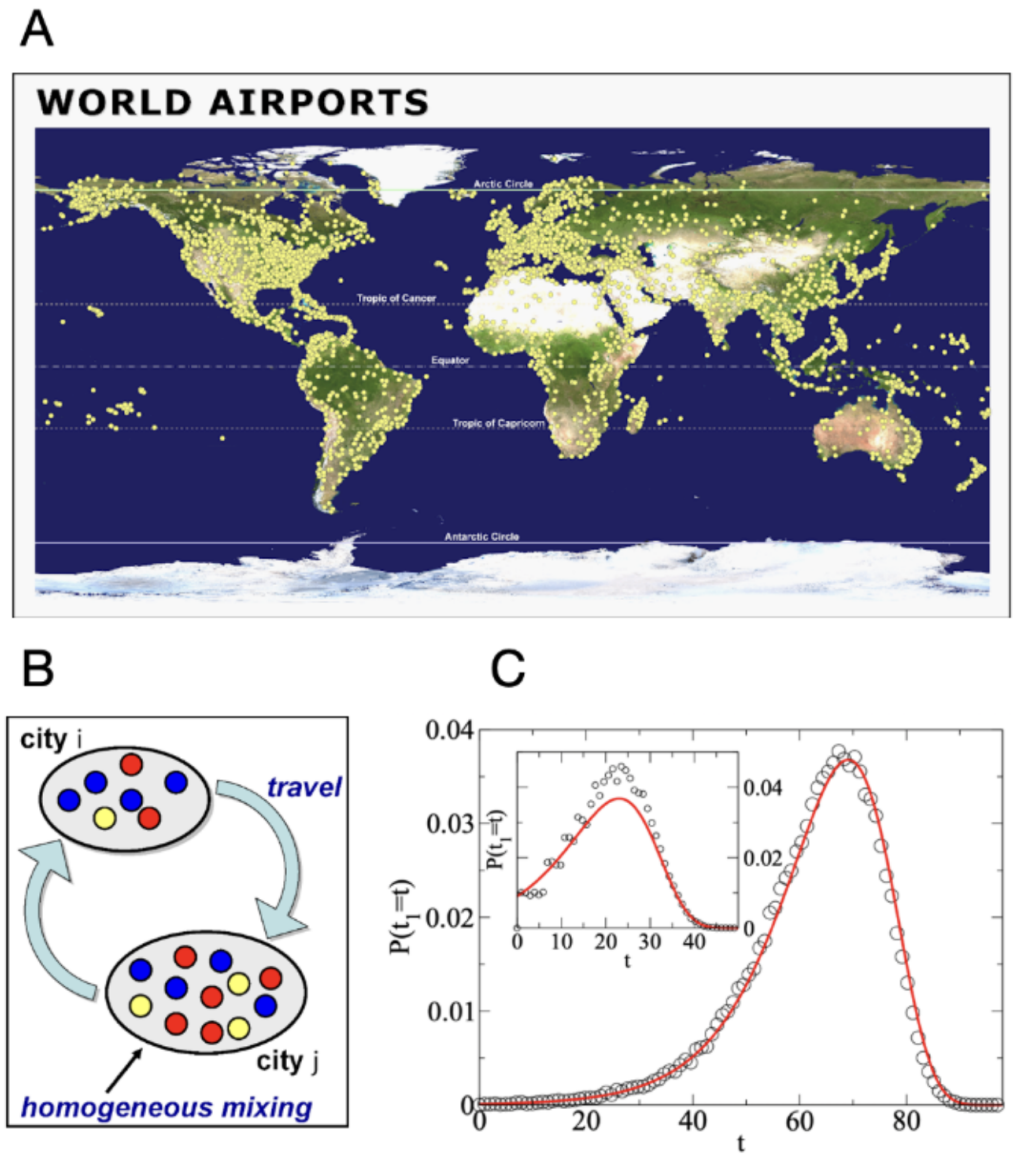


FIGURE 2.1: (A) From [97] Representation of the metapopulation model. Yellow dots represent the airports that constitute the nodes of the network. (B) Illustration of the basic principle of a metapopulation model between two cities, i and j, interconnected by a flow of travellers. (C) From (109). Gumbel distribution underlying the probability of the first arrival time of a new virus in a city 1 connected to a city 0.

to describe better disease specificities [9, 10, 67, 114, 115]. For instance, the model developed in [9, 10, 114] was a SEIR model stratified by age used in France throughout the pandemic. The "Infected" state was divided into four compartments, ranging from least severe to most severe symptoms: asymptomatic, pauci-symptomatic, mild, and severe. A presymptomatic compartment accounted for when an infected individual had not yet exhibited symptoms. Contacts between individuals in the baseline scenario were calculated from contact matrices derived from pre-pandemic surveys [116]. These matrices were then modified to reflect the change in behaviour due to interventions such as school closures, lockdowns, curfews or teleworking. The transmission rate was fitted using hospitalisation cases. Counting hospitalisation was considered more robust than reported cases, as the detection rate did not constrain it. This model

was used to address several public health issues in France. In [114], the authors compared scenarios to exit the first lockdown in May 2020. A few months later, in January 2021, the same model was used to describe the interplay between the emergence of the Alpha variant and the implemented measures to understand the dynamics of cases. They showed that the measures implemented in the first months of 2021 (curfew, closure of large commercial centres, recommendation on teleworking) were strong enough to limit the spread of the historical strain but too weak to contain the emerging Alpha variant. Then, they provided scenarios to compare the impact of strengthening, maintaining, or relaxing social distancing measures [10]. They predicted that a mild lockdown could prevent some regions from facing a third wave. The model was also adapted to understand better the heterogeneity at the regional level [117]. In particular, they showed that the third lockdown was as effective as the second one. They also estimated that a curfew starting at 6 pm was much more effective than a curfew beginning at 8 pm. A similar framework was used to evaluate the impact of the first lockdown and the level of immunity afterwards [14] and inform policymakers of the best way to use lockdown as a last resort [118]. In [119], the authors used a compartmental model to retrospectively study whether regional lockdowns targeting areas with higher incidence would have been more efficient than a national lockdown. The stratification by age was beneficial to assess the effectiveness of age-specific measures. In [120], the authors showed that shielding elderlies was insufficient to relax social distancing measures, as there is an important infection flow between age groups. The model was used to assess the different prioritisation strategies for vaccines in the early roll-out stage by adding compartments for vaccinated people. In [76], the authors showed that age was the most crucial factor to consider in prioritisation. In the case where vaccination only impacted the severity of symptoms, vaccinating elderlies first is more effective in averting hospitalisations and deaths. The effect was found to be weaker if vaccination reduces transmission and susceptibility. To assess the effectiveness of mass testing, the model was also adapted to consider testing and isolation. In [121] infectious people who tested positive were assumed to have a reduced transmission rate to model the effects of self-isolation. This study suggested that widespread testing would be necessary to control a quickly growing epidemic.

2.2.1 Agent-based models.

In the compartmental models described in the previous section, parameters such as the contact rate between age classes [122] are averaged. This framework does not allow for the study of the virus spread at the individual level. Agent-based models (ABM) can account for the heterogeneity of individual interactions [123]. In ABMs, the state of each distinct individual (S, E, I, R, etc.) is explicitly computed at each time step, and interactions with other individuals, which mediate transmission, are modelled at the individual level. Early models assumed a simple contact network between individuals based on regular lattices [124, 125]. These models were later enriched by more realistic contact networks [126]. Some models focused on specific settings such as hospitals or schools. For example, in [127, 128], the authors use face-to-face contact data to construct a temporal contact network within schools. With this framework, they tested different screening strategies (reactive or regular screening, quarantine, symptoms-based isolation) under different vaccination coverage assumptions in the context of the COVID-19 pandemic. Other models simulated the virus spread in a general population across different settings (households, workplaces, schools, public transport…). Before the COVID-19 pandemic, ABMs were used to assess the effectiveness of intervention [129] in various epidemics. With influenza in Southeast Asia,

this framework was used to study targeted mass prophylactic use of antiviral drugs [130], quarantine, and pre-vaccination campaigns [131]. Targeted vaccination was also modelled in the context of smallpox [132], Malaria [133] and RSV [134].

With the COVID-19 pandemic, multilayer agent-based models were particularly useful in evaluating the effectiveness of contact tracing strategies [135, 136], as several NPIs were focused on specific settings (closure of leisure or public places, teleworking enforcement). In [136], the authors used a multi-layer network to account for the diversity of contacts based on locations and household compositions. With this network, they tested scenarios with different assumptions about testing, case tracing, and social distancing measures. In [59], the population was age-stratified to study the impact of digital contact tracing, as several parameters vary with age: contact rate, susceptibility, transmissibility, and smartphone penetration rate in the population.

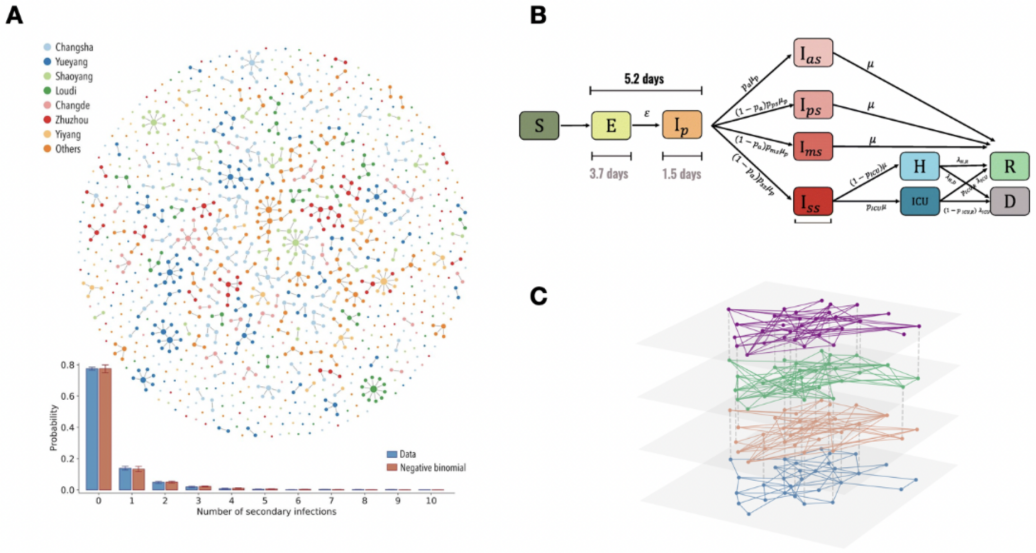


FIGURE 2.2: Representation of various modelling approaches for studying the propagation of SARS-CoV-2 and assessing the impact of interventions. (A) : From [137] Reconstructed chains among first detected cases in Hunan Province. (B) : From [114] Adapted SEIR model to take into account the different phases of the disease. (C) : Structure of an Agent-Based Model with a multi-layer network.

2.3 Conclusion.

In this chapter, I reviewed the approaches employed during the COVID-19 pandemic to model the early global spread of the virus and its variants and to quantify the effectiveness of interventions. I dedicated particular attention to mechanistic models which, unlike purely statistical frameworks, aim at describing the mechanisms driving virus transmission. This approach requires knowledge and data from various fields: mathematics, statistics, biology, epidemiology, public health and immunology [138]. Models synthesise different scales, from intra-host dynamics to spatial propagation. While these models are always a simplified version of reality, they allow the formulation of hypotheses so that they can be tested. For both studying the dissemination of a virus and assessing the effectiveness of an intervention, it is difficult to disentangle the relative roles of different drivers. Mathematical models can test a set of parameters to quantify the relative role of different ingredients and to provide an understanding of how these ingredients interact. In the case of global invasion, different aspects can

affect the international dissemination of a virus. Travelling, local mixing, and sociodemographic features interact to drive the propagation. This was particularly true with the emergence of new COVID-19 variants. In that period, countries were implementing different levels of interventions resulting in a large heterogeneity in local transmission. When assessing the effect of interventions, disentangling the effect of the global epidemic context and the synergy with other interventions is often challenging. The effectiveness of public health interventions on a population scale cannot be assessed by statistical protocols comparing individuals such as random control trials [139]. Mathematical models make it possible to test different scenarios in silico and to compute counterfactual scenarios. However, it is important to assess the limits of the model and to account for uncertainties that mainly come from three different sources: stochastic uncertainties, parameter and model uncertainties [140]. Stochastic uncertainty comes from the intrinsic stochasticity of the studied phenomena - as virus transmission - or data generation - as sampling detected cases. This type of uncertainty can be informed by the proper use of confidence intervals. Sensitivity analysis allows the comparison of scenarios obtained with a wide set of parameters to compute how the findings of a study are robust to inputs. Model uncertainty refers to the assumptions underlying any modelling framework. These assumptions must be clearly discussed and stated. Overall, it is critical to communicate carefully about the possible outcome predicted by a model and to be transparent about the model's limitations to support decision-making.

Chapter 3

Drivers and impact of the early silent invasion of SARS-CoV-2 Alpha.

In this chapter, I present the work published in [5] where I analyse the drivers and impact of the silent spread of the Alpha variant. I first present in detail the data used in this analysis and what is the purpose of this work. Then, I will present the article.

3.1 Introduction.

Originally launched in 2008 to facilitate the sharing of influenza genome sequences, GISAID [43] has evolved to support the rapid open-access sharing of sequences of viruses implied in various diseases, such as influenza, COVID-19, or monkeypox. For Sars-Cov-2, genomic surveillance was critical, notably to identify and study new VOCs that raised important public health concerns. The use of GISAID has become pivotal in the analysis of genomic sequences: in the first two years of the pandemic, 78% of high-income countries sequenced more than 0.5% of their cases and around 25% of these genomes were submitted within 21 days to the platform [141].

The New and Emerging Respiratory Virus Threats Advisory Group (NERVTAG) in the UK identified the Alpha variant as more transmissible in December 2020. This variant was the first to prompt an international alert and to be declared a VOC on December 18, 2020 [142]. However, retrospective analysis of genomic sequences in GISAID shows that Alpha sequences were collected in the three months before the alert, suggesting low attention during three months. During the fall of 2020, most countries implemented mitigation measures to prevent a second wave of cases. The UK entered a lockdown on November 5th. Due to this epidemiological context, the spreading potential was highly heterogeneous in the UK and across European countries and varied in time.

In our study, we used metadata of sequences submitted to GISAID between August 15, 2020, and June 1, 2021. We extracted three metadata: the country where the sequences were collected, the date of sample collection (collection date) and the submission date to the GISAID platform. By analysing these data, we estimated the date of the first detection of the Alpha variant for each country. We defined this date as the collection date of the first Alpha sequence submitted to the GISAID platform in that country. In theory, if all cases were detected and sequenced in time, these data could allow us to reconstruct the spread of the Alpha variant. As shown in Figure 3.1 the Alpha variant was first detected in early September 2020 in the United Kingdom, then

in Denmark and Portugal in late October. However, these data should be considered in perspective with other evidence.

FIGURE 3.1: Collection date of first submitted sequence of Alpha variant in each country on GISAID.

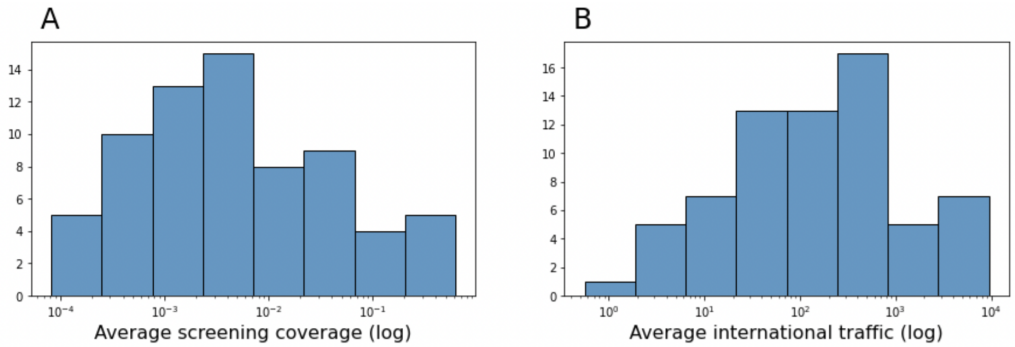


FIGURE 3.2: Histogram of sequencing coverage (A) and International traffics (B) averaged from 1st September 2020 to 31st December 2020.

First, the time lag between the first detection in the UK and in other countries is quite long (over two months) and does not align with studies predicting the exportation of the Alpha variant based on international travel data [34, 84] or phylogenetic studies [38]. Furthermore, virological surveys suggest that, for many countries, the first detection of the Alpha variant likely occurred when local transmission was already installed. Indeed, in France, the first case was detected on December 25, 2020, but a flash survey conducted three weeks later [47] suggested that 3% of observed cases were due to the Alpha variant. To reconcile observations, we investigated the influence of different factors on the propagation and detection of the Alpha variant using various data sources.

Some insight can be obtained by simple theory extending the calculations of Chapter 2, where we reported the computation of the average time of first arrival in a

country of a new virus or variant. The formulas showed that this duration was proportional to the logarithm of the probability of travelling:

$$<t_{\text{first introduction}}> = -\frac{1}{a} ln(\frac{p}{a}).$$

We can extend this formula to the time to detection of first virus or variant in a country. If we note s the probability for variant to be detected once it is imported, then the mean time of first detection is given by

$$<t_{\text{first detection}}> = -\frac{1}{a} ln(\frac{s.p}{a}).$$

In the following article, we analyse the relationship between the time of first detection of the Alpha variant, the probability of travelling and the sequencing coverage. We show that the data about these three parameters follow the relationship derived just before. To estimate bias introduced by the heterogeneity in sequencing between countries, we defined a sequencing coverage index that measures the strength of the surveillance system. This index was computed as the ratio between the number of samples sequenced at a given day, and the number of COVID-19 cases reported the same day. We also measured the time lag between the collection date and the sequence submission date. Additionally, we used data on the number of passengers air travelling from the UK to each studied country [143]. Raw analysis of these data showed significant heterogeneity between countries (see Figure 3.2). Then, we used a Bayesian model to integrate this data and reconstruct the invasion of the Alpha variant, from its emergence and growth in the UK to its exportation, detection, and submission to GISAID in other countries. This approach allowed us to estimate the silent spread for each country, defined as the time between the first introduction and the first detection of the Alpha variant.

To validate model estimates on the date of first importation, we simulated local spread following importation in six destination countries using two different approaches (a branching process [8] and a two-strain compartmental model [10]). We found that our reconstruction is coherent with early virological studies in countries other than the source. Finally, we used the branching process to study the local spread of the Alpha variant.

3.2 Article: Drivers and impact of the early silent invasion of SARS-CoV-2 Alpha.

nature communications

Article

https://doi.org/10.1038/s41467-024-46345-1

Drivers and impact of the early silent invasion of SARS-CoV-2 Alpha

Received: 2 October 2023

Accepted: 22 February 2024

Published online: 09 March 2024

Check for updates

Benjamin Faucher[1], Chiara E. Sabbatini[1], Peter Czuppon[2], Moritz U. G. Kraemer[3,4], Philippe Lemey[5], Vittoria Colizza[1,6,9], François Blanquart[7,9], Pierre-Yves Boëlle[1,9] & Chiara Poletto[8,9]

SARS-CoV-2 variants of concern (VOCs) circulated cryptically before being identified as a threat, delaying interventions. Here we studied the drivers of such silent spread and its epidemic impact to inform future response planning. We focused on Alpha spread out of the UK. We integrated spatio-temporal records of international mobility, local epidemic growth and genomic surveillance into a Bayesian framework to reconstruct the first three months after Alpha emergence. We found that silent circulation lasted from days to months and decreased with the logarithm of sequencing coverage. Social restrictions in some countries likely delayed the establishment of local transmission, mitigating the negative consequences of late detection. Revisiting the initial spread of Alpha supports local mitigation at the destination in case of emerging events.

In December 2020, one year after SARS-CoV-2 emergence, the increased transmissibility and severity of the Alpha variant (Pango lineage B.1.1.7) prompted an international alert[1,2]. Attempts to contain the variant in the UK, where it was first identified, were too late and its global dissemination led to a resurgence of cases and deaths in many countries. Sequences shared through GISAID[3] in real time provided records of the variant's international spread[4] and a number of studies predicted the first countries that would be invaded based on international travel from the UK[5–7]. Still, observations were not in agreement with the expectations, and it soon became clear that the first Alpha detection in countries outside the UK occurred when the variant had been circulating silently in these territories for some time. For instance, the first case infected by the Alpha variant was identified on 25 Dec 2020 in France[3]; yet, three weeks later, already 3% of the ~100,000 weekly reported COVID-19 cases were caused by the Alpha lineage[8]. Late detection was also noted in Switzerland[9] and the USA[10,11].

Phylodynamics analysis and modeling studies revealed that silent spread occurred for early SARS-CoV-2 lineages and subsequent variants of concern (VOCs)[12–20]. This has sparked a public health debate. The efforts to contain a variant at the source are ineffective if they come too late, when the virus is already spreading cryptically out of the source. Interventions aiming at mitigation or delay may instead have an impact depending on the extent and duration of silent dissemination at the time they are implemented[21,22]. Recent works addressed the minimal sequencing coverage to detect a variant early enough for an effective response, and proposed modeling tools for risk assessment[23–27]. However, the complex interplay of factors determining the duration of silent propagation remains poorly understood. Indeed, SARS-CoV-2 VOCs emerged in a context of changing patterns of genomic surveillance, international travel, population immunity, and local interventions. When Alpha emerged in late 2020, sequencing coverage was highly variable and changed dramatically as countries increased genomic surveillance. It took months from the emergence to declaring Alpha a VOC[2]. During this period the epidemiological context across many regions changed substantially. The efforts in the UK and other countries to control a substantial autumn pandemic wave

[1]Sorbonne Université, INSERM, Institut Pierre Louis d'Epidémiologie et de Santé Publique (IPLESP), F75012 Paris, France. [2]Institute for Evolution and Biodiversity, University of Münster, Münster 48149, Germany. [3]Department of Biology, University of Oxford, Oxford, UK. [4]Pandemic Sciences Institute, University of Oxford, Oxford, UK. [5]Department of Microbiology, Immunology and Transplantation, Rega Institute, Laboratory for Clinical and Epidemiological Virology, KU Leuven, Leuven, Belgium. [6]Department of Biology, Georgetown University, Washington, DC, USA. [7]Center for Interdisciplinary Research in Biology, CNRS, Collège de France, PSL Research University, Paris 75005, France. [8]Department of Molecular Medicine, University of Padova, 35121 Padova, Italy. [9]These authors contributed equally: Vittoria Colizza, François Blanquart, Pierre-Yves Boëlle, Chiara Poletto. e-mail: chiara.poletto@unipd.it

impacted the rate of exportations of Alpha out of the UK and the chance to seed local transmission. This makes the emergence of Alpha a paradigmatic example.

Here we used a Bayesian model to retrospectively reconstruct the initial international dissemination of Alpha from 20 Sep 2020 to 31 Dec 2020 out of the UK. By leveraging diverse sources of data for the temporal and geographical change in international travel, sequencing coverage and local epidemic growth, we show that these factors, together with the effect of the international VOC alert on surveillance, drove the duration of Alpha silent spread.

Results

Factors contributing to the spread of Alpha

The early spread of the Alpha variant in the UK occurred in the last quarter of 2020, in a context where a lockdown in the UK, from 5 Nov to 2 Dec 2020, reduced local transmission and the potential for international propagation[28–30]. Air, train, Channel Tunnel and ferry passengers traveling out of the UK in this month had fallen up to 20% of that in September (Fig. 1A).

Over the same period, more than 200,000 sequences were submitted to GISAID from 73 countries, which allowed monitoring the spread of Alpha. We defined the date of first Alpha detection in each country as the date of collection of the first Alpha sequence submitted to GISAID. We hypothesized that sequences collected earlier but submitted at a later date resulted from retrospective surveillance and would misrepresent the routine screening effort. Sequencing coverage ranged over four orders of magnitude over countries: 59% of the cases reported in New Zealand over Sep–Dec 2020 were sequenced, but the median for all countries was only at 0.3%. As might be expected, the date of first detection of Alpha was earlier with higher sequencing coverage and more travelers from the UK (Fig. 1B). The UK was the only country to report the Alpha strain before Dec 1, 2020, followed by Denmark (2 Dec 2020) and Australia (7 Dec 2020). The Alpha international alert on 18 Dec 2020, led to a rise in sequencing coverage (Fig. 1C), shorter collection-to-submission times for Alpha sequences than for others (27 days (CI [8,137]) vs. 52 days (CI [10,162]), Fig. 1D and Supplementary Fig. 1) and prioritization of sequencing of travelers from the UK[4,31]. Nineteen countries collected their first Alpha sequence the week following the alert and submitted it with a median delay of 9 days. In most of these countries, the first case detected was a case imported from the UK[32].

We developed the Alpha international dissemination model to fit the date of first detection and the corresponding date of submission between the beginning of September and end of December in the 73 countries contributing to GISAID during the period. We used dates for 24 countries where the Alpha was detected during the period (including the UK) and accounted for no detection in the other countries by statistical censoring. The key assumption of the model is that the hazard of submitting an Alpha sequence in a country outside the UK results from the dynamically changing incidence in the UK, outbound flows of travelers from the UK, sequencing coverage at arrival and the delay from collection to submission. Thus, we assumed that before the end of December, the first detected cases were traveling cases[4,32] and dissemination was at its early stage, i.e. traveling cases were traveling out of the UK. Although a simplification, this is in line with earlier work showing that the UK was the main source of Alpha dissemination during the first three months, while other countries became more important at a later stage[19]. Time-varying incoming travelers from the UK, sequencing coverage and collection-submission delays were derived from data for each country. Fitted parameters were the exponential growth rate in the UK before and after the beginning of the November lockdown and the increase in genomic surveillance among travelers compared to cases in the community in destination countries following the international alert. Details are given in the Methods section.

Observed dates of first detection and submission (Fig. 2A) and a cumulative number of countries submitting an Alpha sequence (Fig. 2B) matched the model predictions. Portugal and Germany detected Alpha earlier than predicted by our model; there the delays from collection to submission were the longest (48 days for Portugal and 23 days for Germany, versus a median of 9 days in the other countries submitting Alpha). For Portugal, the long gap between the collection dates of the first and the second submitted sequences suggests a retrospective investigation. The model predicted a median seeding date of the Alpha epidemic in the UK on 8 Sep 2020 (95% prediction interval [Aug 21, Sep 19])[33]. The estimated doubling time of incidence in the UK was 4.2 days (95% crI [3.6, 5.3]) before 5 Nov 2020 and 10.6 days (95% crI [6.5, 22]) afterwards. Assuming the reproductive ratio $R = 1 + rT$, with T the generation time interval at 6.5 days[34] and r the Alpha exponential growth in the UK, these estimates would be compatible with $R = 2.0$ [1.8, 2.3] and $R = 1.4$ [1.1, 1.65] before and after 5 Nov 2020. These values broadly agree with previous estimates, with a pattern of decreased transmission over time[28–30,33,35]. With these estimates, the predicted trend of Alpha infections in the UK was in agreement with the observations (Fig. 2C)[36]. The large number of countries reporting Alpha almost simultaneously in late December was explained by an estimated 50-fold (95% crl [12, 298]) increase of sequencing coverage among travelers compared to non-travel related cases following the alert, consistently with the active search of Alpha cases among travelers and their contacts. Further details on parameter estimates and fit convergence are reported in Supplementary Fig. 2 and Supplementary Table 5.

In a sensitivity analysis, results were found to be robust to a range of modeling assumptions—e.g. changepoints for the exponential growth of incidence in the UK, rate of detection of COVID-19 infections outside the UK, and incubation period among the others. Details are reported in Supplementary Table 5.

Silent spread ranged from days to weeks

We next used the international dissemination model to predict the date of the first introduction of Alpha from the UK to each of the 73 countries under study and the duration of silent spread, i.e. the duration of the time from the first introduction to the first detection of Alpha. We found that up to ~65 countries could have experienced the introduction of Alpha by the end of December, compared with the 24 countries that reported it (Fig. 3A). Our model predicted that the first introduction of Alpha in a country occurred up to 70 days earlier than the date of first Alpha detection (Fig. 3B, C). For instance, our model predicted that Alpha arrived 60 days earlier in Italy with an average sequencing coverage of 0.3% during the period, while it was only 15 days in Hong Kong with a sequencing coverage of 50%. Overall, the duration of the silent spread showed a logarithmic association with the average sequencing coverage (Fig. 3D). The estimated dissemination pattern is consistent with real-time projections based on air-travel[5]. Early introductions in Denmark and the USA were also consistent with the result of phylodynamic analyses and retrospective surveillance[10,11,37–39]. We found that the collection date of the first Alpha sample ever collected in each country (earlier than the first detection in 34 countries because of retrospective surveillance) was within the range of first introduction predicted by the model but for Colombia.

Local dynamics affected the impact of silent spread

We then focused on the spread of Alpha in six countries where national genomic investigations estimated the incidence of the Alpha variant in early January 2021: Denmark, France, Germany, Portugal, Switzerland, and the USA. We used a stochastic model (autochthonous model A)[40] to simulate chains of transmission generated by infections introduced from the UK as predicted by the international dissemination model described above. The model used

https://doi.org/10.1038/s41467-024-46345-1

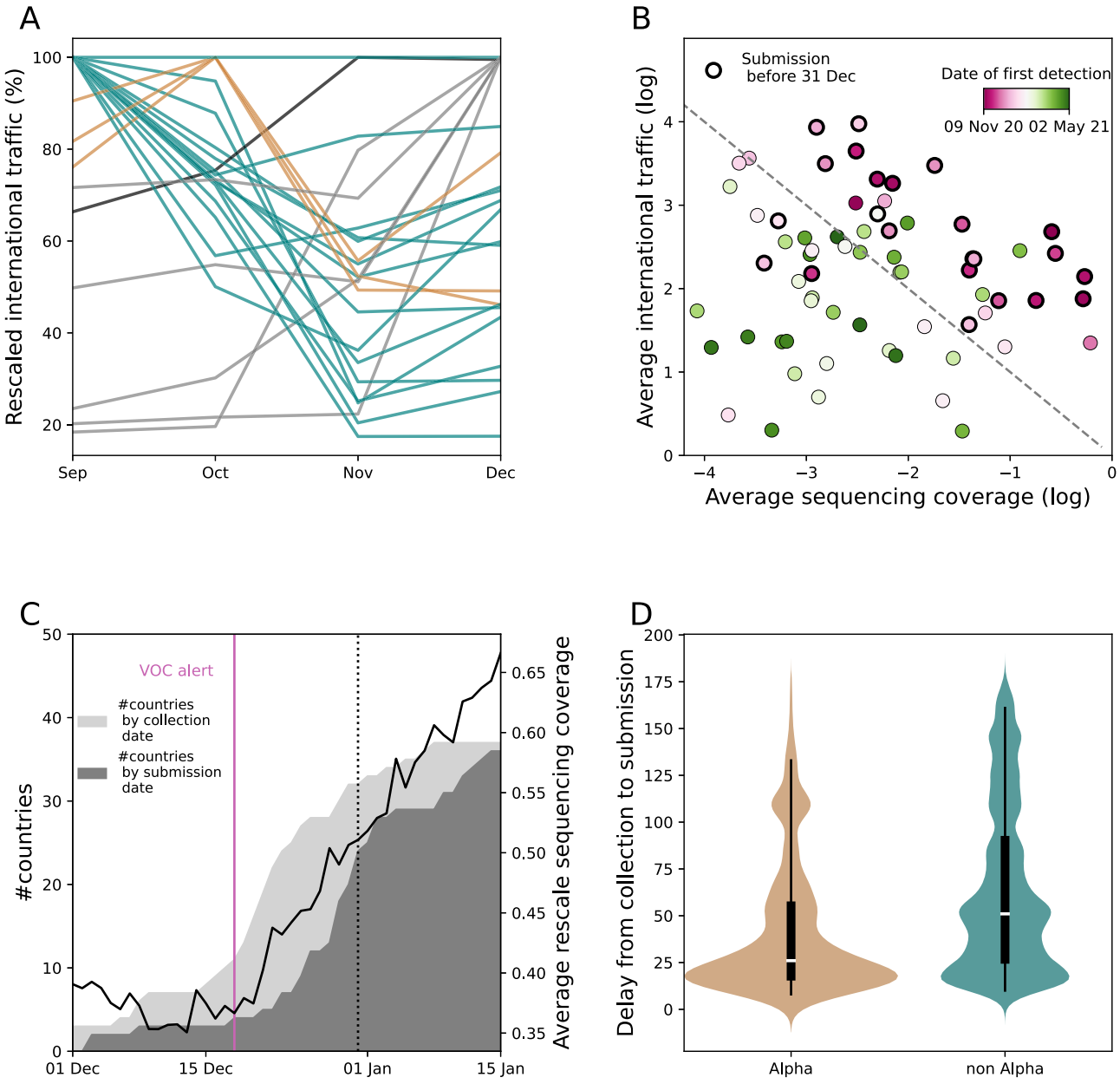


Fig. 1 | Factors associated with the pattern of observed Alpha dissemination. **A** Change in outbound international traffic from the UK over time, including air-travel, train, ferry and Channel Tunnel[59]. The 73 countries contributing to GISAID during 1 Sep 2020–31 Dec 2020 are shown as an example. Traffic is rescaled to the maximum over the period. To improve readability, different months of traffic maximum are associated with a different color. **B** Date of first detection, i.e. collection of the first Alpha sequence submitted to GISAID, for each of the 73 countries as in (**A**), according to sequencing coverage and international traffic (passengers/day) averaged over 1 Sep 2020–31 Dec 2020. For each day, the sequencing coverage of a country is defined as the number of collected SARS-CoV-2 sequences on GISAID −regardless the date of submission−divided by reported cases. The dashed line provides a guide to the eye, as, under simplifying assumptions[44,81], we expect the date of first detection to be a function of log(sequencing coverage) + log(traveling flaw) (Supplementary Information). **C** Number of countries with at least one Alpha submission plotted by date of collection and date of submission. The black line shows the average rescaled sequencing coverage. In each country, the sequencing coverage was rescaled by the maximum over the period displayed in the plot to highlight the trend. Countries' rescaled time series were then averaged. For the sake of visualization, the sequencing coverage is here smoothed over a 2 weeks sliding window. The purple line indicates the date of Alpha international alert (18 Dec 2020). The dashed black line indicates the censoring date used in the analysis (31 Dec 2020). **D** Distributions of delay (in days) from collection to submission for Alpha and non-Alpha sequences collected outside the UK from December 2020 to mid-January 2021 and submitted up to June 2021 (non-Alpha sequences $n = 149699$, Alpha sequences $n = 6992$). Boxplots represent the median (white bar), the quartiles and the 95% range (whiskers). The violin plot shows the Kernel estimation of the underlying distribution. Additional details are reported in Supplementary Fig. 1.

country-specific time-varying reproduction number, overdispersion in transmission, and a 60% transmission advantage of Alpha over the wildtype[28,29,41]. The model reproduced the same trend of the observed Alpha cases with a case ascertainment fraction around 50% (Fig. 4A). Incidence in the USA was underestimated, possibly due to heterogeneity in the different states. To test a finer spatial resolution we retrieved Alpha frequency data for California, Florida, and New York City, obtaining a good match with the data for California and New York City and an under-estimation (within the range of possible stochastic outcomes) for Florida (Supplementary Fig. 4). To test the robustness of these predictions, we used a second model with age-structure, temporal variation in social contacts due to restrictions,

https://doi.org/10.1038/s41467-024-46345-1

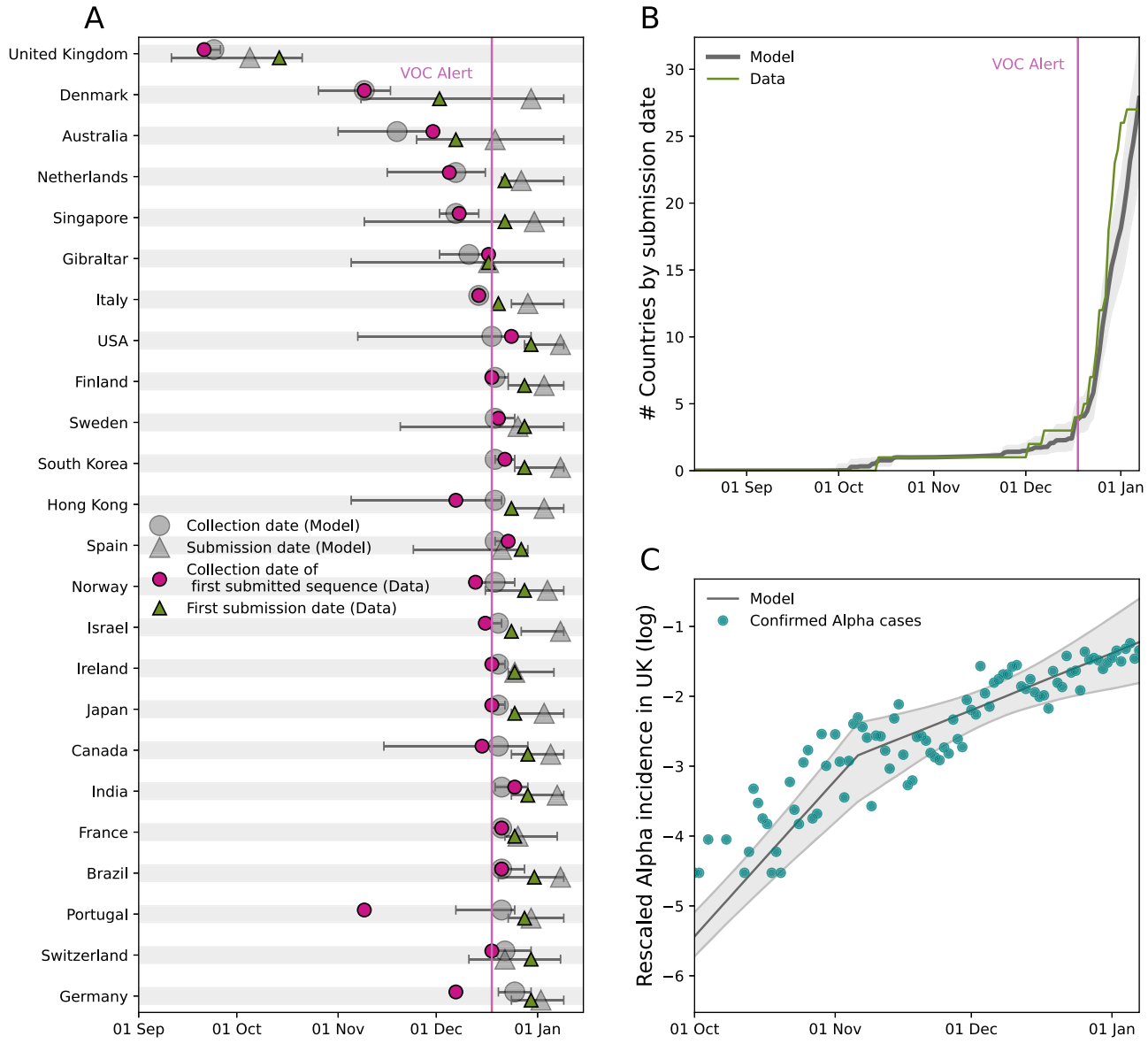


Fig. 2 | Comparison between the international dissemination model and the data. A Date of collection of the first Alpha sample submitted to GISAID and corresponding date of submission for the 24 countries submitting Alpha sequences before 31 Dec 2020. Data are shown by purple circles (collection) and green triangles (submission). Median date obtained from the model is indicated by gray circles (collection) and gray triangles (submission). The horizontal bars display the 95% prediction interval over n = 500 simulations. **B** Median model predicted cumulative number of countries submitting a first Alpha sequence to GISAID compared with observations. In panels A and B, the purple vertical line indicates the date of Alpha international alert (18 Dec 2020). **C** Alpha incidence in the UK[36] and median model-predicted epidemic profile in the UK. Both model predictions and data are rescaled to the sum over the period considered to allow comparing the profiles of the curves. To account for testing delays model predictions are shifted right of one week. The gray colored ribbon represents the 95% credible interval.

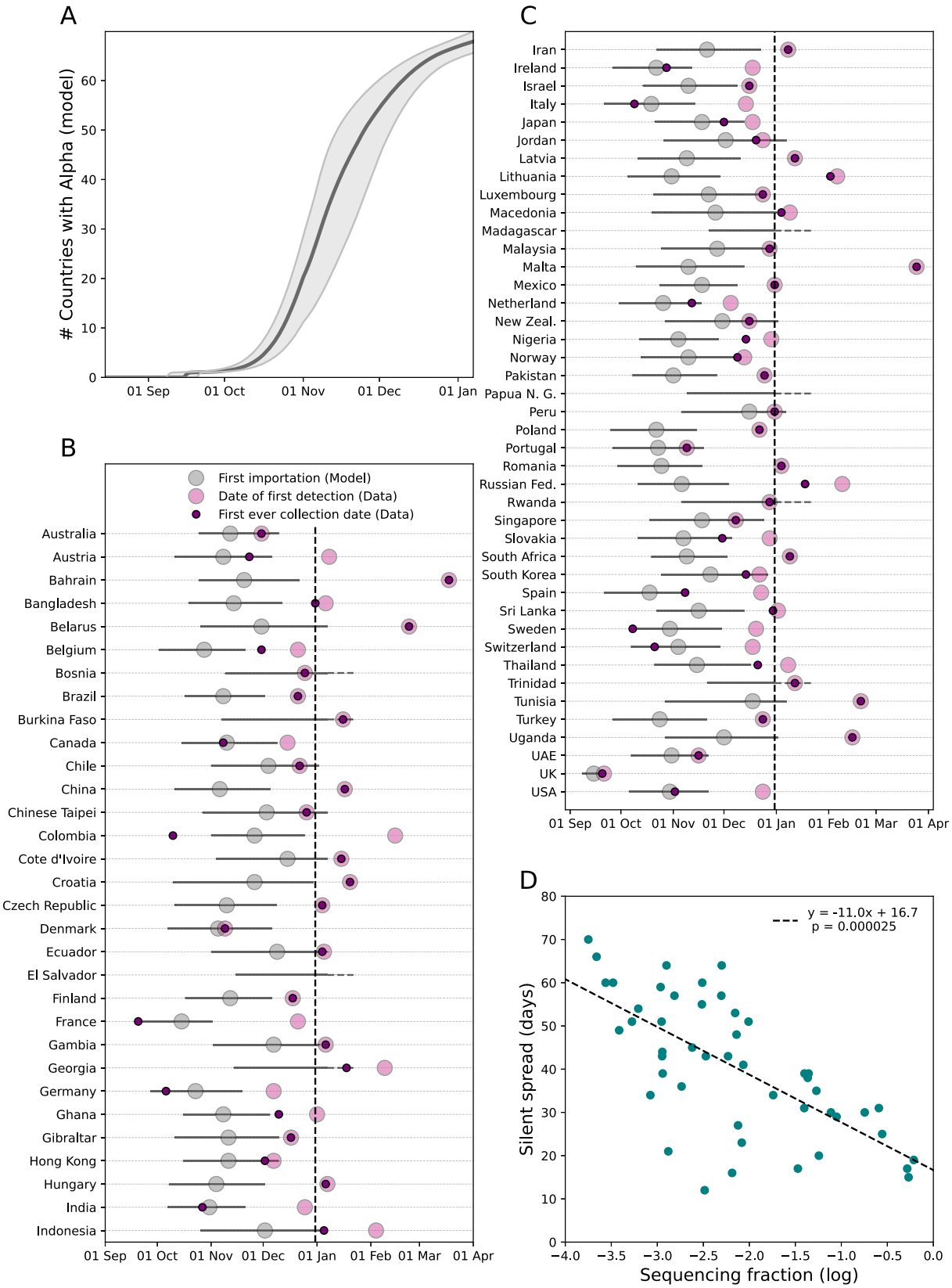
and the co-circulation between Alpha and wildtype that was calibrated and validated for France[42,43] (autochthonous model B) finding also in this case a good agreement (inset of Fig. 4A).

Besides supporting our estimates of Alpha dissemination out of the UK, the reconstruction of local epidemics outside the UK allowed investigating the potential impact of silent spread in the six focal countries. The estimated Alpha cases as of 31 Dec 2020 broadly scaled with the international traffic connecting the country with the UK, showing the important role of importations in determining local Alpha epidemic size (Fig. 4B). Still, potential consequences of silent spread could only be gauged by taking into account changes in local transmission (Fig. 4C). For example, while the first detected case in Germany and Switzerland had been collected with a similar delay from the predicted date of first importation, the reproductive ratio R_t in Germany had generally been larger than in Switzerland during the period. Therefore, the seeding of transmission chains still active at the end of the year in Germany could take place well before the first detected case was collected for the first time in the territory, while in Switzerland ~50% of the transmission chains started after first virus detection (Fig. 4C). Overall, later seeding of active chains was associated with smaller average R_t over the period (Fig. 4D), but not with the reduction in traveling (Supplementary Fig. 5). Therefore, our analysis suggests that low levels of local R_t enhanced the relative contribution of late importations, potentially countering the consequences of late detection.

Discussion

Genomic surveillance has been a major advancement in monitoring the spread of SARS-CoV-2 after initial emergence. However, interpreting these data is complicated as they do not follow a pre-

https://doi.org/10.1038/s41467-024-46345-1



established and coordinated sampling design. Retrospective analyses of the past dissemination of VOCs can provide epidemiological knowledge that enables us to better respond to future viral emergence events. Here, taking the initial Alpha spread as an example, we showed that several components of the highly heterogeneous epidemic context had to be taken into account for interpretation.

Previous studies focused on traveling flows to explain the arrival of a first infection into a new country[44–48]. Yet differences in genomic surveillance capacity, over four orders of magnitude across countries during Alpha emergence, profoundly affected the introduction-to-detection delay with a logarithmic decrease in sequencing coverage. Furthermore, extraneous events like the international alert further

Fig. 3 | Timing of first importation and silent spread as estimated by the international dissemination model. A Cumulative number of countries with an Alpha introduction as predicted by the model. The quantity is computed from the median predicted date of introduction in each country with 95% prediction interval obtained over n = 500 simulations. **B**, **C** Median date of first introduction when occurring before Dec 31 (vertical dashed line) for each country estimated by the model with 95% prediction interval computed over $n = 500$ simulations. For each country, we report the date of first Alpha detection (i.e. collection of first submitted sequence) (light pink) and the date of the first ever collected Alpha sequence (dark pink) from the data. For El Salvador, Papua New Guinea and Madagascar, no Alpha sequence had been reported before June 2021. **D** Duration of silent spread in days vs sequencing coverage. The distribution of the durations of silent spread is reported in Supplementary Fig. 3. Duration of silent spread is computed as the difference between the median date of first detection and the median date of first introduction as predicted by the model. We restricted the analysis to countries for which both first introduction and first detection were predicted to occur before 7 Jan 2021. Dashed line represents least-squares linear regression. *P*-value is computed from Wald test with t-distribution.

altered the speed of variant detection. These strong spatiotemporal changes in genomic surveillance partially masked the true pattern of Alpha invasion, to the point that the correlation between the dates of detection and the international traffic was poor in the first 24 countries reporting Alpha (spearman correlation 0.24, $p = 0.3$). Yet the good fit obtained with the international dissemination model and the increase of Alpha epidemic size with traveling flow (Fig. 4B) both suggest that traveling flows were a driver of viral spread, in agreement with other works[19,44–47,49]. A more uniform sequencing collection protocol would have provided a coherent view of Alpha propagation improving public health awareness and response. This highlights the importance of eliminating surveillance blind spots by increasing sequencing in countries with poor surveillance[23].

According to our model, Alpha was introduced in more than 60 countries before the international alert. This provides evidence that when an emerging pathogen is not reported in a given destination country, it may likely be due to the surveillance system not yet being able to detect it. The alert triggered heightened genomic surveillance worldwide, reinstated lockdown measures in the UK, and resulted in border screening and travel bans in countries connected to the UK[23,28–30]. However, international response arrived at a moment in which Alpha was already widespread in several countries, preventing containment. Improving surveillance across countries would reduce the time from importation to detection, but it would still clash with the delay needed to recognize a novel variant as a VOC. A lineage with important mutations can be identified relatively quickly if sequencing coverage is high enough[23,24,27], although the assessment of clinical risk is slower[24]. Lineages have shown the ability to become dominant without any increase in fitness in particular epidemiological contexts[50], while others like Beta remained at low frequency despite mutations of clinical importance. A more rapid recognition of Alpha as a VOC could have advanced the response by health authorities to delay the establishment of Alpha during a time when vaccination became available in some countries[21]. Similar delays in declaring a VOC were also observed for subsequent VOC episodes[19]. This underlines the complexity of the interpretation of a context with emerging new variants[51]—especially when major known drivers such as international travel are in place—and of the decision-making for public health response.

The growing Alpha epidemic in the UK allowed dissemination despite the drop in international traffic out of the UK and the social restrictions in many countries. For instance, while UK travelers to France dropped by 56% in November compared to September, the number of Alpha-infected travelers to France still grew from 1 to 10 daily over November 2020 according to our model. The lockdown implemented in France at this time likely did not prevent local transmission because Alpha was more transmissible. Local restrictions may however delay successful invasion, as was apparent from the in-depth analysis of the six destination countries: a lower local reproductive ratio delayed the seeding of local transmission chains following importations up to one month. Although with the same analysis we could not address the consequences of the decline in travel, we expect that when local transmission is limited by control measures, introductions from the country of origin contribute more substantially to the epidemic at destination[20]. We can thus hypothesize that limiting importation early could act synergistically with local restrictions to limit the size of the VOC epidemic. Still, we expect that the fine tuning between different factors (e.g. quality and extent of borders control and timing of their implementation[22,52,53]) can affect the impact of travel restrictions.

Following Alpha, other SARS-CoV-2 variants raised concern due to their rapid emergence and spread, namely Beta, Gamma, Delta, Omicron and its sublineages. Undetected introductions and silent spread were likely common to all variants, although the epidemic context progressively changed between 2021 and 2022. The rise in international mobility and social contacts accelerated the spread of Delta and Omicron[19]. This has reduced the window for public health response requiring an intensification of virus genomic surveillance to enable authorities to identify variants in time. However the high costs of genomic surveillance and the phasing out of the pandemic have now reduced our ability to detect future VOC emergence events. The Alpha experience shows the importance of designing sequencing protocols able to balance sustainability and detection capacity by meeting the minimal requirements of sequencing extent and reporting delay—e.g. sequencing 0.5% of cases with a turnaround time smaller than 21 days as previously proposed[23], and by leveraging information from multiple sources, including wastewater and animal surveillance[54,55]. Importantly, this study also highlights that the knowledge of surveillance extent and protocol adopted by countries is key to real-time data analysis to better assist risk assessment and intervention planning. This would be facilitated by the widespread adoption of pre-established surveillance protocols.

Our study is affected by a number of limitations. First, sequencing coverage was computed at the country level and no distinction could be made for traveler vs. local cases due to the poor available information on testing rate among travelers[18]. We dealt with this by allowing an increase in detection after the Alpha alert. Second, we analyzed here the period before 31 Dec 2020. This time window was long enough to cover the seeding from the UK to the destination countries and observe the consequent onset of local transmission. At the same time, the window is sufficiently short to assume in first approximation the UK to be the source of Alpha spread, before large epidemics in other countries became the dominant source of traveling cases. Extending the analysis to a longer period would require a more general framework that can be the subject of future work. Third, it is not possible to set a cut-off between real-time and retrospective surveillance when computing sequencing coverage from GISAID metadata. The computation of sequencing coverage being affected by retrospective surveillance could potentially overestimate the extent of the real-time genomic surveillance. Fourth, we have here defined the date of first Alpha detection in a country as the date of collection of the first sequence submitted to GISAID. Reporting of variants of interest to local public health authorities can be indeed more rapid than submitting the sequence to GISAID. Still, we acknowledge that this may depend on the country and stage of the invasion, e.g. before and after the alert. In addition, the public sharing of a variant's sequence enables the recognition of its presence in a given territory by a larger public, including health authorities and the scientific community worldwide.

https://doi.org/10.1038/s41467-024-46345-1

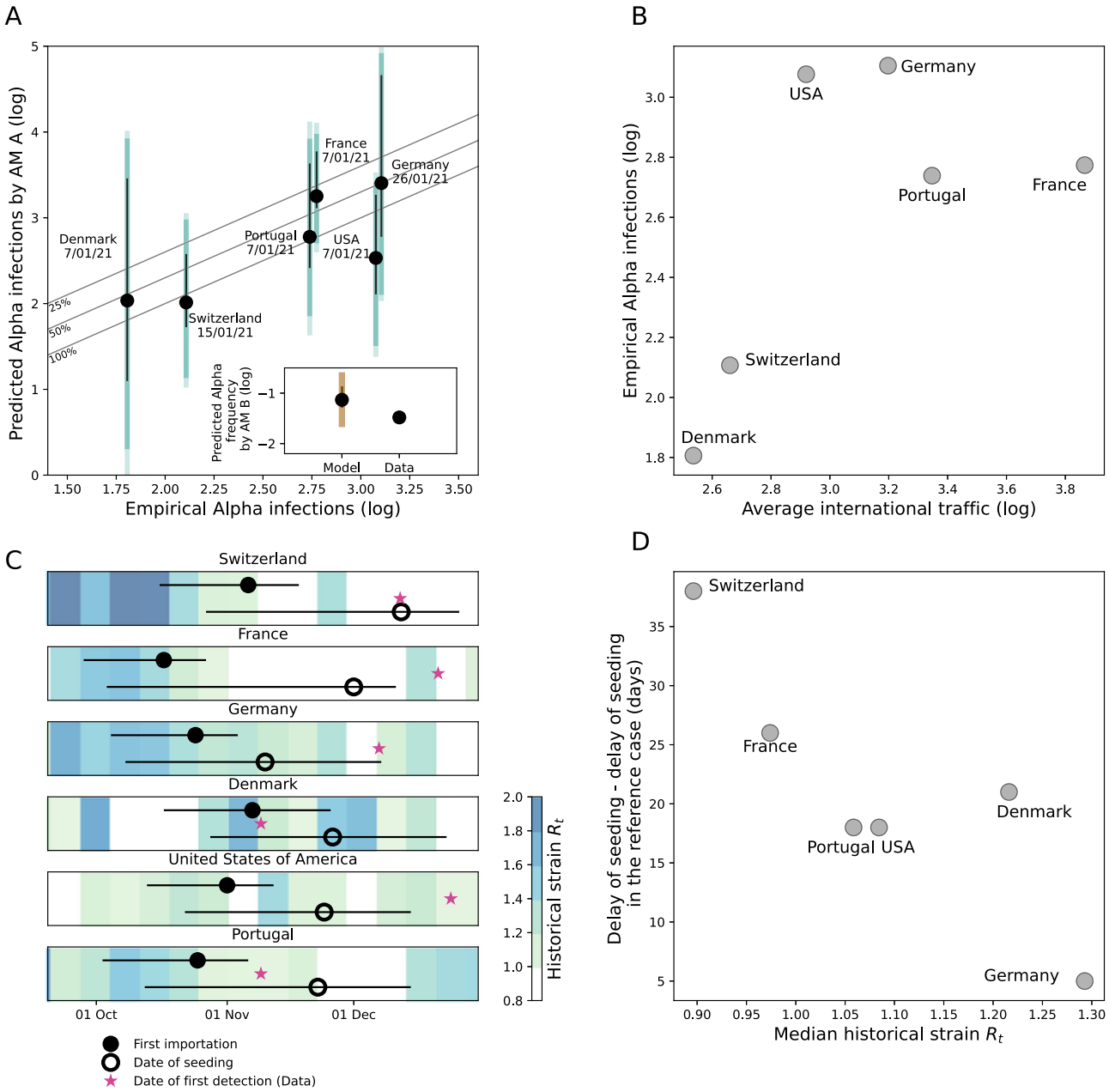


Fig. 4 | Local spread of Alpha in six destination countries. A Model vs. empirical Alpha infections. In the main plot, the empirical estimates of Alpha cases are computed by multiplying the Alpha frequency from virological investigations by the reported COVID-19 incidence at the same date—the date is indicated in the plot. Model estimates are obtained with the autochthonous model A (AM A in the plot). Gray lines show ratios of 100%, 50% and 25% between observed and predicted infections attributable to reporting. In the inset, the frequency of Alpha in France obtained from the autochthonous model B (AM B in the plot) is compared with the empirical data. In both panels, black error bars indicate the prediction interval over 500 stochastic simulations obtained with the median volume of Alpha introduction, output of the international dissemination model assuming a 7-day delay between case and infection. Dark colored bars account for the variability in the output of the autochthonous models accounting for the upper and the lower limit of the prediction interval of the Alpha introductions as given by the international dissemination model. Light colored bars account for variability in the delay from infection to case reporting (ranging from 4 days to 10 days). **B** Empirical Alpha infections vs average international traffic. **C** Comparison between the date of first introduction as predicted by the international dissemination model and the seeding time of the transmission chains survived until 31 Dec 2020, predicted with the autochthonous model A. Circles indicate medians and segment the 95% prediction interval. Colors indicate the effective reproduction number of the historical strain, R_t, computed from weekly mortality data (Methods). The star shows the date of first Alpha detection as a comparison. **D** Difference between the median delay of seeding predicted by the autochthonous model A and the same quantity in the reference case—i.e. when R_t is the same in all countries and traveling fluxes do not change in time, plotted against the median R_t during the period from first introduction to seeding.

To conclude, by jointly modeling epidemic dissemination and observation based on GISAID submissions we have quantified Alpha silent spread in countries outside the UK unveiling its link with international travel and sequencing coverage. Our results show that the duration of Alpha silent spread varied from days to months. Strong spatiotemporal heterogeneities in surveillance provided a major obstacle to data interpretation. Still, restrictions in place in destination countries may have delayed the establishment of local transmission and partially mitigated the negative consequences of late detection and response. By the time a new variant is recognized as a potential

threat, surveillance authorities of countries outside the variant source should be prepared for the variant being potentially already present in the territory. Enhancement in local screening and measures aiming at containing local transmission are thus key ingredients of a response plan. Taken together these findings provide lessons learnt for the future management of SARS-CoV-2 variants. Beyond that, retrospective reconstructions of SARS-CoV-2 spread are essential to improve computational modeling and public health knowledge to better guide the response to future spread of viruses with zoonotic and pandemic potential.

Methods

Data

GISAID records. While we did not use any actual sequences in this study, from GISAID entries[3] we retrieved collection dates, submission dates, information on lineage (i.e. whether it was Alpha or not) and country for all human SARS-CoV2 sequences submitted between 15 Aug 2020 and 1 Jun 2021 included (n = 1,735,675 downloaded on 2 Jun 2021). Data in GISAID originated from 144 countries, however, only 73 countries had submitted sequences collected between September 2020 and December 2020. We used GISAID entries to determine the date of the first submission of an Alpha sample in each country and the respective date of collection, the latter defined as the "detection date". Assuming that detection occurs at the time of sample collection corresponds to the optimistic hypothesis that surveillance authorities are informed right after a sample is collected. We also computed the date of the first collection ever of an Alpha sample in each country, irrespective of the date submitted. Finally, we determined the distribution of delays from collection to submission and the sequencing coverage from the number of sequences by country and date of collection (see below). For GISAID sequences missing a collection date (3%), we imputed the missing date with a date selected at random from the sequences with complete data submitted in the same week and country. We resorted to imputation instead of inferring a statistical model because of the small percentage of missing records. In addition, the strong spatiotemporal variations displayed by the data could be hard to capture by a statistical model.

COVID-19 cases and death data. We retrieved the daily number of COVID-19 cases by country from the COVID-19 data repository hosted by the Center for Systems Science and Engineering at Johns Hopkins University (CSSE)[56] to compute the sequencing coverage. Incidence of Alpha cases in the UK was obtained from the "Variants of Concern: technical briefing 7—Data England" report[36]. We used the weekly deaths time series from the European Center for Disease Prevention and Control[57] (downloaded on 1 Jul 2021) to compute the time varying reproduction ratio in Denmark, France, Germany, Portugal, Switzerland and the USA.

Travel data. Travel flow from the UK to destination countries was reconstructed combining air travel data, estimates of passengers via train, Channel Tunnel and ferries. We computed probabilities of travel assuming a catchment population of 36 M for London airports. More precisely:

- Air travel data were obtained from the International Air Transport Association (IATA)[58]. It comprised the monthly number of passengers outbound from English airports by country of destination. From the monthly data we computed an averaged daily flux of passengers over the month. For each country, we aggregated all passengers directed to the country and leaving from all airports of London.
- Eurostar rail passenger numbers going each day to France, Belgium and the Netherlands were estimated as in[59], assuming a 95% reduction due to the COVID-19 pandemic[60].
- We used the monthly number of cars crossing the Channel Tunnel[59,61] to derive an averaged daily flux of passengers over the month. We assumed that 1.5 passengers travel on average for each car[59] and that the repartition of passengers among countries in continental Europe is the same as for trains.
- Numbers of passengers via ferries to France, Belgium, the Netherlands, Spain and Ireland were obtained by ref. 62. We used monthly data to compute an averaged daily flux of passengers over the month.

Virological investigation records. National investigations were conducted in a number of countries in early January. Through bibliographic search and via social media we gathered the data from virological surveys or extensive screening for Denmark, France, Germany, Portugal, Switzerland, the USA. These surveys give an estimated frequency of Alpha infections for the cases detected a given day (or a given time period). We also considered the daily number of detected cases on the day of the survey (or the midpoint of the time period) from CSSE. From these two numbers, we calculated the number of detected Alpha cases. In Supplementary Table 1 we report the source, the date of the survey, detected Alpha frequency, and the number of Alpha cases computed for each country. In Supplementary Fig. 4, we also analyzed three locations in the US, i.e. Florida, California and New York City. Sources for these data are reported in Supplementary Table 2.

Data processing

Sequencing coverage. The sequencing coverage was computed for each day and each country as the number of sequences collected after imputation divided by the number of cases. In Fig. 1C, we smoothed the sequencing coverage with a two-week sliding window to highlight the general trend.

Delays from collection to submission. We computed the collection-to-submission times in different ways before and after the Alpha alert on 18 Dec 2020. Before the alert, we hypothesized that Alpha sequences would be reported with the same time pattern as other sequences. We therefore computed a delay distribution by country and by date of collection using all GISAID sequences as $\pi_c(d;u) = \frac{n_{u+d,c}}{N_{u,c}}$ where d is the delay, $N_{u,c}$ the number of sequences collected on day u in country c and $n_{u+d,c}$ those submitted on date $u+d$. For sequences collected after the alert of 18 Dec 2020, we accounted for the different delay distribution for Alpha and other sequences. Due to the limited number of Alpha sequences collected outside the UK soon after the alert we aggregated all data collected outside the UK, thus defining an average Alpha delay distribution for all countries. We then used a 3-day smoothing time window, where length 3 was chosen as the best compromise to smooth out fluctuations without masking meaningful trends. We therefore computed $\pi_c(d;u) = \frac{n_{d+u}}{N_u}$ with N_u the number of Alpha sequences collected between day $u-1$ and $u+1$, and n_{u+d} the number of those sequences submitted after d days for each country c. Delays from collection to submission are reported in Fig. 1D and Supplementary Fig. 1. In the sensitivity analysis we computed the Alpha collection-to-submission delays after 18 Dec 2020 separately for each country. We used a longer smoothing time window (7 days instead of 3 days) to compensate for the geographic disaggregation.

International dissemination model

We model the observed data consisting in date pairs $\{S_c, T_c\}$ by country, where S_c is the date of first submission of an Alpha sequence to GISAID and T_c the corresponding date of collection in country c. The model is based on the following assumptions: i) Alpha incidence in the UK grows exponentially with a piecewise exponential rate to account for the autumn lockdown; ii) imported cases are proportional to international traffic; iii) collection and sequencing of a sample from an

imported case of SARS-CoV-2 Alpha and its submission on GISAID is proportional to sequencing coverage and the detection-to-submission delay computed from GISAID metadata.

More precisely, we first described incident Alpha infections in the UK at time t as exponentially growing with time according to $inc_{UK}(t) = \exp(\sum_{T_0}^{t} r(u))$, where T_O is fixed at 15 Aug 2020[33], the date when the risk of emergence starts and $r(t)$ the daily exponential growth rate. The daily exponential growth rate in the UK was considered piecewise constant, r_1 up to Nov 5th, 2020, when the UK entered a lockdown, and r_2 afterwards. In other words, $inc_{UK}(t)$ was a "two-slope" exponential, growing as $\exp(r_1 t)$ before Nov 5 and as $\exp(r_2 t)$ afterwards. We also explored a model with no change of slope and two changes of slopes (at 5 Nov 2020 and at 2 Dec 2020, beginning and end of the lockdown respectively) in the sensitivity analysis.

In the UK, the number of Alpha sequences collected depended on incidence and sequencing coverage as

$$\lambda^*_{UK}(t) = K_{UK} s_{UK}(t) \sum_{j=0}^{J} inc_{UK}(t-j), \quad (1)$$

where $s_{UK}(t)$ is the sequencing coverage on day t, J the duration of incubation and K_{UK} the detection probability. For the incubation period we used 5 days[63] and tested 4 and 6 days in the sensitivity analysis. We considered that one case out of 4 would be tested ($K_{UK} = 0.25$)[64].

Consistently with[4,32] we assumed that the first case reported to GISAID in each country outside the UK was an imported case, infected in the UK but discovered abroad. Thus, we modeled detection and sequencing in countries outside the UK without the need to model local variant growth. There, the expected number of sequences collected at time t in country c additionally accounted for traveling as

$$\lambda^*_c(t) = K_c \, p_c(t)/N s_c(t) \sum_{j=0}^{J} inc_{UK}(t-j), \quad (2)$$

where $p_c(t)/N$ is the fraction of the population traveling from the catchment area of the London airports to country c on day t with $N = 36$ millions inhabitants the population of the area, and $s_c(t)$ the sequencing coverage in country c on day t and K_c the fraction of imported infections being detected as COVID-19 cases. We assumed detection of imported cases to be higher than the detection of local cases, thus we used $K_c = 0.5$ ($> K_{UK}$). In the sensitivity analysis, we tested all airports of England, instead of airports of London, as the origin of Alpha infected travelers, and $K_c = 0.25$. Finally, we allowed for an increase in collection of Alpha sequences among travelers relative to others after the alert of 18 Dec 2020 due to increasing sampling of travelers from the UK[4,32] using a multiplicative factor γ. Therefore, the expected number of collected Alpha sequences on day t is $\lambda_c(t) = \lambda^*_c(t)$ before 18 Dec 2020 and $\lambda_c(t) = \gamma \lambda^*_c(t)$ afterwards. Taking into account collection-to-submission time, the expected number of sequences submitted at time t in country c is therefore $\alpha_c(t) = \sum_{u \le t} \lambda_c(u) \pi_c(t-u, u)$, and the probability that a sequence submitted on day t was collected on day u, with $u \le t$, is $\lambda_c(u)\pi_c(t-u,u)/\alpha_c(t)$.

To write up the likelihood of observations, we considered that the model described the dynamics of collection and submission until the end of 2020. We assumed Poisson variability in the number of Alpha infections and computed the probability that an Alpha sequence is submitted on GISAID for the first time on date S_c in country c as

$$P(S_c) = \exp\left(-\sum_{u<S_c} \alpha_c(u)\right)\left(1 - \exp(-\alpha_c(S_c))\right) \quad (3)$$

The log-likelihood of the data in the model was:

$$\log L(\{r_0, r_1\}, \gamma; \{S_c, T_c\}) =$$

$$= \sum_{c: S_c \le D} \log(1 - \exp(-\alpha_c(S_c))) + \log(\lambda_c(T_c)\pi_c(S_c - T_c, T_c)/\alpha_c(S_c)) - \sum_c \sum_{T_0}^{S_c} \alpha_c(u), \quad (4)$$

where the first sum runs on countries where an Alpha sequences was submitted before date D (= 31/12/2020) and the second runs in all countries. The summary of all fixed parameters and their values is reported in Supplementary Table 3.

The model likelihood was explored with a Metropolis-Hastings procedure using R v4.3. We used an Exp(0.1) exponential prior on the first exponential growth rate r_1, a N(0,1) prior on second growth rate r_2 to allow for negative growth and an Exp(0.01) prior for the increase in sampling γ (Supplementary Table 4). Unless stated otherwise, 3 chains were run in parallel for 100000 iterations, with the first 50000 discarded as burn-in, the second half was thinned (1 iteration every 25) for a final posterior sample of size 2000. Convergence of the chains was checked visually (Supplementary Fig. 2). Estimates and credible intervals for the fitted parameters are reported in Supplementary Table 5 (baseline values, first row).

We computed the predictive distribution for the date of detection given the actual travel and sequencing coverage as

$$F_c(t; p_c, s_c, K_c) = 1 - \exp\left(-\int_{T_0}^{t} \lambda_c(u; p_c, s_c) du\right) \quad (5)$$

using the posterior sample and computed 95% prediction intervals from these samples.

We finally computed the model-predicted date of first introduction in country c as the distribution $F_c(t; p_c, 1, 1)$ in each country, assuming that 100% sequencing occurred ($s = 1$) and all cases were detected ($K = 1$).

We computed predictive distributions from the model using parameters taken in the posterior distribution as follows (where the "hat" notation corresponds to the estimated value):

- Expected incidence in the UK:

$$inc_{UK}(t) = \exp\left(\sum_{T_0}^{t} \hat{r}(u)\right) \quad (6)$$

- Distribution of time of emergence in the UK:

$$P(T_e < t | T_e < T_{UK}) = 1 - \exp\left(-\sum_{T_o}^{t} \hat{r}(u)\right) \Big/ \left(1 - \exp\left(-\sum_{T_o}^{T_{UK}} \hat{r}(u)\right)\right) \quad (7)$$

- Cumulated distribution of date of first submission:

$$P(S_c \le t) = 1 - \exp\left(-\sum_{u \le t} \hat{\alpha}_c(u)\right) \quad (8)$$

- Cumulated distribution of date of first introduction:

$$P(I_c \le t) = 1 - \exp\left(-\sum_{u \le t} \hat{\lambda}^1_c(u)\right) \quad (9)$$

with

$$\lambda_c^1(t) = p_c(t)/N \sum_{j=0}^{J} inc_{UK}(t-j) \quad (10)$$

the number of (detected and undetected) infections.

To visualize goodness of fit, we computed the cumulated number of countries submitting an Alpha sequence by date t as $\sum_c P(S_c \le t)$, and for the countries reporting an Alpha sequence, the cumulative distribution of introduction date conditional on submission date, $P(I_c \le t | S_c)$.

Autochthonous model A

To simulate the number of Alpha variant infections at the beginning of 2021 in each country of interest, we used the daily rates of importation as estimated from the international dissemination model $\lambda_c^1(t)$) and simulated the subsequent stochastic outcome of each imported infectious individual in the destination country. The different Alpha epidemic clusters initiated by each importation were assumed to be independent. The stochastic epidemic growth model has been described elsewhere[40]. For each day since T_0 and each country of destination, we drew the number of imported infections in a Poisson distribution with rate $\lambda^1{}_c(t)$. Then, starting with each imported infection, we simulated an epidemic chain assuming that each infected individual produced a number of secondary infections according to a negative binomial distribution with mean $(1+\alpha)R_t$ and dispersion parameter $\kappa = 0.4$, where R_t is the effective reproduction number at date t and $\alpha = 0.6$ is the transmission advantage of the Alpha variant relative to the historical strain, assumed to be the same in every country[41]. The generation time distribution was gamma with mean 6.5 days and s.d. 4 days (shape 2.64, scale 2.46)[29]. To compute the effective reproduction number R_t of the historical strain from mortality data, we computed first the daily exponential growth rate as $r_t = 1/7 \log(D_{w+1}/D_w)$ where D_w is the number of deaths in week w. To account for the lag between disease onset and death (approx. 3 weeks), we considered that this exponential growth rate applied to infections for days t in week $w-3$. We finally computed $R_t = \int_0^\infty \exp(-r_t\tau) g(\tau) d\tau$ with $g(\tau)$ the generation interval distribution[65]. Note that the calculation of R_t in this way is robust to under reporting biases, provided that the reporting ratio does not change substantially over the period. This approach yielded estimates similar to the Epiestim method[66].

The model was implemented in C++ (v11). In the simulations of epidemic clusters, the code loops over time, starting from one infected individual at the day of importation, and ending at 31 Jan 2021. Time was discretized in time-steps of 0.1 day. The secondary infections are added to their (future) date in the incidence table, and the code proceeds to the next infected individual at this time step, then to the next time-step. Five hundreds (500) replicate simulations were obtained for each country to account for stochastic variability in the number and timing of importations and growth of local epidemics.

Number of infections output of the model were compared to the empirical number cases estimated from the virological survey. Assuming a delay between infection and case detection of one week, empirical cases were compared with model-predicted Alpha infections 7 days before. Since delay in reporting may vary from one country to another—some countries report cases by date of testing, others by date of notification, data may be smoothed, etc.—we also tested delays of 4 and 10 days.

Autochthonous model B

We used a stochastic discrete age-stratified, two-strain transmission model to simulate the epidemic dynamics in France generated by the estimated Alpha importations[42,43,67].

The model integrates data on demography, age profile, social contacts, mobility and adoption of preventive measures. Four age classes are considered: [0–11), [11–19), [19–65) and 65+ years old (children, adolescents, adults and seniors respectively). Transmission dynamics follows a compartmental scheme specific for COVID-19 where individuals are divided into susceptible, exposed, infectious, hospitalized and recovered. The infectious class is further divided into prodromal, asymptomatic and symptomatic. Susceptibility and transmissibility depend on age[68–70]. Transmissibility also depends on the level of symptoms[71–74].

Contact matrices are setting-specific. Contacts at school are modeled according to the French school calendar, while those at work and on transports according to the workplace presence estimated by Google data[75]. During the different stages of the pandemic, physical contacts are modulated based on surveys on the adoption of physical distancing[76], self-protection[42], and assuming a reduction in contacts due to severe symptoms. The integration of all these data allows for capturing the social distancing restrictions put in place in France to curb the second wave, namely a lockdown with schools open[77] from week 44 (starting October 31, 2020) to week 51 (ending December 15, 2020).

The model was previously used to respond to the COVID-19 pandemic in France in 2020[42,43,63,78], assessing the impact of lockdown[63], of night curfew[43] and of the reopening of schools[78], estimating the underdetection of cases[42], and anticipating the impact of the Alpha variant in France[43]. In particular, we used, here, the same two-strain version of the model developed to study the impact of January 2021 curfew in France on the Alpha circulation in the territory[43], with same parametrization and same transmissibility calibrated to national daily hospital admission data[79]. This accounts for the co-circulation of Alpha variant and the historical strains, and assumes complete cross-immunity between the two strains, higher hospitalization rate and an increase in transmissibility of 50% for Alpha[28]. We also tested a 60% advantage in transmission, finding that results were robust. Values of other key parameters are generation time equal to 6.6 days, and incubation period 5.2 days. Other parameter values are reported in ref. 63. The model was implemented in Python 3.8.5.

We simulate the epidemic dynamics using the output of the international dissemination model as seeding for the dynamics. At each date, we extract the number of prodromal adults infected with the variant from a Poisson distribution with mean equal to the traveling cases at that date obtained from the international dissemination model. We repeat this extraction for each of the 500 stochastic runs performed and we simulate the resulting outbreak. We then compute the proportion of Alpha on January 8 and compare it with the proportion identified by the first large-scale genome sequencing initiative (called Flash #1)[41] conducted in the country on January 7-8, 2021 (Alpha proportion in France equal to 3.3%).

Seeding time of active transmission chains

The time of seeding of a transmission chain still active at a reference end time (time T_R) is uniformly distributed over the range of possible introduction times when the exponential growth rate r is the same in the place of origin (here the UK) and in the destination country and traveling flows are constant over time. This is because starting from the date of emergence T_E, the number of introductions in the destination country at some time t_I will be proportional to $\exp[r(t_I - T_E)]$ and each case introduced will cause $\exp[r(T_R - t_I)]$ cases at time T_R, so that the overall number of cases at time T_R is $\exp[r(t_I - T_E)] \exp[r(T_R - t_I)] = \exp[r(T_R - T_E)]$ irrespective of the actual date of introduction. Therefore, date $(T_E + T_R)/2$ is the expected median introduction date in this simple scenario of constant exponential growth rate and traveling.

We therefore used the autochthonous model A to reconstruct the distribution of the seeding times for the transmission chains still active

on December 31st, 2020. We computed the distribution of seeding times and the difference between the median of this distribution and the expected median under the constant exponential growth rate and traveling described above. The extent of this difference illustrates the effect of the actual change in epidemic growth rate and traveling flows on seeding success. We are here interested on how this quantity changed across the six countries. We found that it increased for lowering values of R_t. This show that low values of R_t were likely hindering the seeding of local transmission chains by the introduced cases, making the late importations comparatively more important.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The findings of this study are based on metadata associated with a total of 1,735,675 sequences available on GISAID and submitted between 15 Aug 2020 and 1 Jun 2021 included and downloaded on 2 Jun 2021 via gisaid.org (GISAID: EPI_SET_230724tv). To view the contributors of each sequence associated with the metadata we used, visit https://doi.org/10.55876/gis8.230724tv. Proprietary airline data are commercially available from OAG and IATA databases (https://www.iata.org/). All other data used in the study are publicly available online. Channel Tunnel data were obtained from https://www.eurotunnelfreight.com/fr/2021/01/trafic-navettes-du-mois-de-decembre-2020/, ferries data were obtained from https://www.gov.uk/government/statistical-data-sets/sea-passenger-statistics-spas, COVID-19 cases were obtained from https://github.com/CSSEGISandData/COVID-19, COVID-19 deaths were obtained from https://www.ecdc.europa.eu/en/publications-data/data-national-14-day-notification-rate-covid-19, Alpha cases in the UK were obtained from https://assets.publishing.service.gov.uk/media/6059e4ad8fa8f545d5c339fc/Variants_of_Concern_VOC_Technical_Briefing_7_England.pdf.

Code availability

Source codes to reproduce the results of this study are publicly shared on zenodo[80] and on github (https://github.com/EPIcx-lab/COVID-19/tree/master/Adherence_and_sustainability).

References

1. Preliminary genomic characterisation of an emergent SARS-CoV-2 lineage in the UK defined by a novel set of spike mutations. *Virological* https://virological.org/t/preliminary-genomic-characterisation-of-an-emergent-sars-cov-2-lineage-in-the-uk-defined-by-a-novel-set-of-spike-mutations/563 (2020).
2. NERVTAG: *Meeting on SARS-CoV-2 variant under investigation VUI-202012/01*, December 2020. *GOV.UK* https://www.gov.uk/government/publications/nervtag-meeting-on-sars-cov-2-variant-under-investigation-vui-20201201-18-december-2020.
3. Shu, Y. & McCauley, J. GISAID: Global initiative on sharing all influenza data—from vision to reality. *Eurosurveillance* **22**, 30494 (2017).
4. O'Toole, Á. et al. Tracking the international spread of SARS-CoV-2 lineages B.1.1.7 and B.1.351/501Y-V2 with grinch. *Wellcome Open Res.* **6**, 121 (2021).
5. Du, Z. et al. Risk for international importations of variant SARS-CoV-2 originating in the United Kingdom. *Emerg. Infect. Dis.* **27**, 1527–1529 (2021).
6. Chinazzi, M. et al. *Preliminary Estimates of the International Spreading Risk Associated with the SARS-CoV-2 VUI 202012/01*. https://www.esdiario.com/images/carpeta_gestor/archivos/2021/01/26/covid_19_uk_new_strain__3__1.pdf.
7. Lai, S., Floyd, J. & Tatem, A. *Preliminary risk analysis of the international spread of new COVID-19 variants, lineage B.1.1.7, B.1.351 and P*. https://hub.worldpop.org/resources/reports/risk_analysis_covid19/Preliminary_risk_analysis_of_the_spread_of_new_variants_24Feb2021.pdf.
8. SPF. *COVID-19: point épidémiologique du 28 janvier* 2021. https://www.santepubliquefrance.fr/maladies-et-traumatismes/maladies-et-infections-respiratoires/infection-a-coronavirus/documents/bulletin-national/covid-19-point-epidemiologique-du-28-janvier-2021.
9. Chen, C. et al. Quantification of the spread of SARS-CoV-2 variant B.1.1.7 in Switzerland. *Epidemics* **37**, 100480 (2021).
10. Washington, N. L. et al. Emergence and rapid transmission of SARS-CoV-2 B.1.1.7 in the United States. *Cell* **184**, 2587–2594.e7 (2021).
11. Alpert, T. et al. Early introductions and transmission of SARS-CoV-2 variant B.1.1.7 in the United States. *Cell* **184**, 2595–2604.e13 (2021).
12. Niehus, R., De Salazar, P. M., Taylor, A. R. & Lipsitch, M. Using observational data to quantify bias of traveller-derived COVID-19 prevalence estimates in Wuhan, China. *Lancet Infectious Dis.* **20**, 803–808 (2020).
13. Pinotti, F. et al. Tracing and analysis of 288 early SARS-CoV-2 infections outside China: A modeling study. *PLOS Med.* **17**, e1003193 (2020).
14. Davis, J. T. et al. Cryptic transmission of SARS-CoV-2 and the first COVID-19 wave. *Nature* **600**, 127–132 (2021).
15. Bedford, T. et al. Cryptic transmission of SARS-CoV-2 in Washington state. *Science* **370**, 571–575 (2020).
16. Worobey, M. et al. The emergence of SARS-CoV-2 in Europe and North America. *Science* **370**, 564–570 (2020).
17. Lemey, P. et al. Accommodating individual travel history and unsampled diversity in Bayesian phylogeographic inference of SARS-CoV-2. *Nat. Commun.* **11**, 5110 (2020).
18. McCrone, J. T. et al. Context-specific emergence and growth of the SARS-CoV-2 Delta variant. *Nature* **610**, 154–160 (2022).
19. Tegally, H. et al. Dispersal patterns and influence of air travel during the global expansion of SARS-CoV-2 variants of concern. *Cell* **186**, 3277–3290.e16 (2023).
20. Tsui, J. L.-H. et al. Genomic assessment of invasion dynamics of SARS-CoV-2 Omicron BA.1. *Science* **381**, 336–343 (2023).
21. Kucharski, A. J. et al. Travel measures in the SARS-CoV-2 variant era need clear objectives. *Lancet* **399**, 1367–1369 (2022).
22. Reichmuth, M. L., Hodcroft, E. B. & Althaus, C. L. Importation of Alpha and Delta variants during the SARS-CoV-2 epidemic in Switzerland: Phylogenetic analysis and intervention scenarios. *PLOS Pathogens* **19**, e1011553 (2023).
23. Brito, A. F. et al. Global disparities in SARS-CoV-2 genomic surveillance. *Nat. Commun.* **13**, 7003 (2022).
24. Brett, T. S. & Rohani, P. Containing novel SARS-CoV-2 variants at source is possible with high-intensity sequencing. *PNAS Nexus* **1**, pgac159 (2022).
25. Susswein, Z. et al. Leveraging global genomic sequencing data to estimate local variant dynamics. https://doi.org/10.1101/2023.01.02.23284123 (2023).
26. Klamser, P. P. et al. Enhancing global preparedness during an ongoing pandemic from partial and noisy data. *PNAS Nexus* **2**, pgad192 (2023).
27. Han, A. X. et al. SARS-CoV-2 diagnostic testing rates determine the sensitivity of genomic surveillance programs. *Nat. Genet.* **55**, 26–33 (2023).
28. Davies, N. G. et al. Estimated transmissibility and impact of SARS-CoV-2 lineage B.1.1.7 in England. *Science* **372**, eabg3055 (2021).
29. Volz, E. et al. Assessing transmissibility of SARS-CoV-2 lineage B.1.1.7 in England. *Nature* **593**, 266–269 (2021).
30. Kraemer, M. U. G. et al. Spatiotemporal invasion dynamics of SARS-CoV-2 lineage B.1.1.7 emergence. *Science* **373**, 889–895 (2021).
31. *Sequencing of SARS-CoV-2—first update*. https://www.ecdc.europa.eu/en/publications-data/sequencing-sars-cov-2 (2021).

32. *Risk Assessment: Risk related to spread of new SARS-CoV-2 variants of concern in the EU/EEA*. https://www.ecdc.europa.eu/en/publications-data/covid-19-risk-assessment-spread-new-sars-cov-2-variants-eueea (2020).
33. Hill, V. et al. The origins and molecular evolution of SARS-CoV-2 lineage B.1.1.7 in the UK. *Virus Evolution* **8**, veac080 (2022).
34. Flaxman, S. et al. Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. *Nature* **584**, 257–261 (2020).
35. Lineage-specific growth of SARS-CoV-2 B.1.1.7 during the English national lockdown. *Virological* https://virological.org/t/lineage-specific-growth-of-sars-cov-2-b-1-1-7-during-the-english-national-lockdown/575 (2020).
36. Public Health England. *SARS-CoV-2 variants of concern and variants under investigation in England, Technical briefing 7*.
37. Michaelsen, T. Y. et al. Introduction and transmission of SARS-CoV-2 lineage B.1.1.7, Alpha variant, in Denmark. *Genome Medicine* **14**, 47 (2022).
38. Chaillon, A. & Smith, D. M. Phylogenetic Analyses of Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) B.1.1.7 lineage suggest a single origin followed by multiple exportation events versus convergent evolution. *Clinical Infectious Diseases* **73**, 2314–2317 (2021).
39. Phylogenetic evidence that B.1.1.7 has been circulating in the United States since early- to mid-November. *Virological* https://virological.org/t/phylogenetic-evidence-that-b-1-1-7-has-been-circulating-in-the-united-states-since-early-to-mid-november/598 (2021).
40. Czuppon, P., Schertzer, E., Blanquart, F. & Débarre, F. The stochastic dynamics of early epidemics: probability of establishment, initial growth rate, and infection cluster size at first detection. *J. Roy. Soc. Interface* **18**, 20210575 (2021).
41. Gaymard, A. et al. Early assessment of diffusion and possible expansion of SARS-CoV-2 Lineage 20I/501Y.V1 (B.1.1.7, variant of concern 202012/01) in France, January to March 2021. *Eurosurveillance* **26**, 2100133 (2021).
42. Pullano, G. et al. Underdetection of cases of COVID-19 in France threatens epidemic control. *Nature* **590**, 134–139 (2021).
43. Domenico, L. D., Sabbatini, C. E., Pullano, G., Lévy-Bruhl, D. & Colizza, V. Impact of January 2021 curfew measures on SARS-CoV-2 B.1.1.7 circulation in France. *Eurosurveillance* **26**, 2100272 (2021).
44. Gautreau, A., Barrat, A. & Barthélemy, M. Global disease spread: Statistics and estimation of arrival times. *J. Theor. Biol.* **251**, 509–522 (2008).
45. Poletto, C. et al. Assessing the impact of travel restrictions on international spread of the 2014 West African Ebola epidemic. *Eurosurveillance* **19**, 20936 (2014).
46. Poletto, C., Boëlle, P.-Y. & Colizza, V. Risk of MERS importation and onward transmission: a systematic review and analysis of cases reported to WHO. *BMC Infectious Dis.* **16**, 448 (2016).
47. Scalia Tomba, G. & Wallinga, J. A simple explanation for the low impact of border control as a countermeasure to the spread of an infectious disease. *Math. Biosci.* **214**, 70–72 (2008).
48. Brockmann, D. & Helbing, D. The hidden geometry of complex, network-driven contagion phenomena. *Science* **342**, 1337–1342 (2013).
49. Gozzi, N. et al. Estimating the spreading and dominance of SARS-CoV-2 VOC 202012/01 (lineage B.1.1.7) across Europe. https://doi.org/10.1101/2021.02.22.21252235 (2021).
50. Hodcroft, E. B. et al. Spread of a SARS-CoV-2 variant through Europe in the summer of 2020. *Nature* **595**, 707–712 (2021).
51. Kraemer, M. U. G. et al. Monitoring key epidemiological parameters of SARS-CoV-2 transmission. *Nat. Med.* **27**, 1854–1855 (2021).
52. Vasylyeva, T. I. et al. Introduction and establishment of SARS-CoV-2 gamma variant in New York City in early 2021. *J. Infectious Dis.***226**, 2142–2149 (2022).
53. Chinazzi, M. *et al*. The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. *Science* https://doi.org/10.1126/science.aba9757 (2020).
54. Diamond, M. B. et al. Wastewater surveillance of pathogens can inform public health responses. *Nat. Med.* **28**, 1992–1995 (2022).
55. Subissi, L. et al. An early warning system for emerging SARS-CoV-2 variants. *Nat. Med.* **28**, 1110–1115 (2022).
56. Dong, E., Du, H. & Gardner, L. An interactive web-based dashboard to track COVID-19 in real time. *Lancet Infectious Dis.* **20**, 533–534 (2020).
57. *Data on 14-day notification rate of new COVID-19 cases and deaths*. https://www.ecdc.europa.eu/en/publications-data/data-national-14-day-notification-rate-covid-19 (2023).
58. IATA. https://www.iata.org/en/.
59. du Plessis, L. et al. Establishment and lineage dynamics of the SARS-CoV-2 epidemic in the UK. *Science* **371**, 708–712 (2021).
60. Eurostar left battling for survival following Covid-19 slump. *Breaking Travel News* https://www.breakingtravelnews.com/news/article/eurostar-left-battling-for-survival-following-covid-19-slump/.
61. Eurotunnel—Trafic Navettes du mois de décembre. *Eurotunnel Le Shuttle* 2020 https://www.eurotunnelfreight.com/fr/2021/01/trafic-navettes-du-mois-de-decembre-2020/.
62. *Sea passenger statistics: data tables (SPAS). GOV.UK* https://www.gov.uk/government/statistical-data-sets/sea-passenger-statistics-spas (2023).
63. Di Domenico, L., Pullano, G., Sabbatini, C. E., Boëlle, P.-Y. & Colizza, V. Impact of lockdown on COVID-19 epidemic in Île-de-France and possible exit strategies. *BMC Med.* **18**, 240 (2020).
64. Colman, E., Puspitarani, G. A., Enright, J. & Kao, R. R. Ascertainment rate of SARS-CoV-2 infections from healthcare and community testing in the UK. *J. Theor. Biol.* **558**, 111333 (2023).
65. Wallinga, J. & Lipsitch, M. How generation intervals shape the relationship between growth rates and reproductive numbers. *Proc. Roy. Soc. B: Biol. Sci.* **274**, 599–604 (2006).
66. Cori, A., Ferguson, N. M., Fraser, C. & Cauchemez, S. A New Framework and Software to Estimate Time-varying Reproduction Numbers during Epidemics. *Am. J. Epidemiol.* **178**, 1505–1512 (2013).
67. Di Domenico, L. et al. Adherence and sustainability of interventions informing optimal control against the COVID-19 pandemic. *Commun. Med.* **1**, 1–13 (2021).
68. Viner, R. M. et al. Susceptibility to SARS-CoV-2 infection among children and adolescents compared with adults: a systematic review and meta-analysis. *JAMA Pediatrics* **175**, 143–156 (2021).
69. Thompson, H. A. et al. Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) setting-specific transmission rates: a systematic review and meta-analysis. *Clin. Infectious Dis.* **73**, e754–e764 (2021).
70. Fontanet, A. et al. SARS-CoV-2 infection in schools in a northern French city: a retrospective serological cohort study in an area of high transmission, France, January to April 2020. *Eurosurveillance* **26**, 2001695 (2021).
71. Li, R. et al. Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV-2). *Science* **368**, 489–493 (2020).
72. Qiu, X. et al. The role of asymptomatic and pre-symptomatic infection in SARS-CoV-2 transmission—a living systematic review. *Clin. Microbiol. Infection* **27**, 511–519 (2021).
73. Buitrago-Garcia, D. et al. Occurrence and transmission potential of asymptomatic and presymptomatic SARS-CoV-2 infections: a living systematic review and meta-analysis. *PLoS Med.* **17**, e1003346 (2020).

74. *Contact Settings and Risk for Transmission in 3410 Close Contacts of Patients With COVID-19 in Guangzhou, China: A Prospective Cohort Study: Annals of Internal Medicine*: Vol 173, No 11. https://www.acpjournals.org/doi/full/10.7326/M20-2671.
75. COVID-19 Community Mobility Report. *COVID-19 Community Mobility Report* https://www.google.com/covid19/mobility?hl=fr.
76. *CoviPrev: une enquête pour suivre l'évolution des comportements et de la santé mentale pendant l'épidémie de COVID-19*. https://www.santepubliquefrance.fr/etudes-et-enquetes/coviprev-une-enquete-pour-suivre-l-evolution-des-comportements-et-de-la-sante-mentale-pendant-l-epidemie-de-covid-19.
77. Pullano, G., Domenico, L. D., Sabbatini, C. E. & Colizza, V. *Expected impact of a lockdown with schools in session—France*, Nov 2020.
78. Di Domenico, L., Pullano, G., Sabbatini, C. E., Boëlle, P.-Y. & Colizza, V. Modelling safe protocols for reopening schools during the COVID-19 pandemic in France. *Nat. Commun.* **12**, 1073 (2021).
79. *Données hospitalières relatives à l'épidémie de COVID-19 (SIVIC)—data.gouv.fr*. https://www.data.gouv.fr/fr/datasets/donnees-hospitalieres-relatives-a-lepidemie-de-covid-19/.
80. Faucher, B. et al. Drivers and impact of the early silent invasion of SARS-CoV-2 Alpha. *Zenodo* https://doi.org/10.5281/zenodo.10639236 (2024).
81. Bajardi, P. et al. Human mobility networks, travel restrictions, and the global spread of 2009 H1N1 Pandemic. *PLoS ONE* **6**, e16591 (2011).

Acknowledgements

The authors gratefully acknowledge all data contributors, i.e., the Authors and their Originating laboratories responsible for obtaining the specimens, and their Submitting laboratories for generating the genetic sequence and metadata and sharing via the GISAID Initiative, on which this research is based. We acknowledge financial support by the Municipality of Paris through the programme Emergence(s) to C.P. and B.F.; Cariparo Foundation through the program Starting Package to C.P.; Department of Molecular Medicine through the program SID from BIRD funding to C.P.; the Agence Nationale de la Recherche project DATAREDUX with grant agreement ANR-19-CE46-0008-03 to V.C.; ANRS–Maladies Infectieuses Émergentes project EMERGEN (ANRS0151) to V.C.; EU Horizon 2020 grants MOOD with grant agreement H2020-874850 (publication cataloged as MOOD096) to V.C., C.P., P.Y.B., M.U.G.K., P.L. and RECOVER (H2020- 101003589) to V.C.; the ERC grant EvoComBac (949208) to F.B.; ERC grant ReservoirDOCS (725422) to P.L.; Marie Sklodowska-Curie action (MSCA) grant PolyPath (844369) to P.C.; Institut des Sciences du Calcul et de la Donnée (ISCD).

Author contributions

V.C., F.B., C.P., and P.-Y.B. conceived and designed the study. B.F., P.-Y.B., and C.P. developed the international dissemination model. P.C., and F.B., developed the autochthonous model A. C.E.S. and V.C. developed the autochthonous model B. M.U.G.K. and P.L. critically commented on the model ingredients and assumptions. V.C., F.B., C.P., and P.-Y.B. wrote the original draft. All authors discussed the results, edited the manuscript, and approved its final version.

Competing interests

The authors declare no competing interests

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-024-46345-1.

Correspondence and requests for materials should be addressed to Chiara Poletto.

Peer review information *Nature Communications* thanks Shihao Yang and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

Supplementary Information

Drivers and impact of the early silent invasion of SARS-CoV-2 Alpha

Benjamin Faucher[1], Chiara E. Sabbatini[1], Peter Czuppon[2], Moritz U.G. Kraemer[3,4], Philippe Lemey[5], Vittoria Colizza[†1,7], Francois Blanquart[†6], Pierre-Yves Boëlle[†1], Chiara Poletto[†*8]

[1]Sorbonne Université, INSERM, Institut Pierre Louis d'Epidémiologie et de Santé Publique (IPLESP), F75012, France

[2]Institute for Evolution and Biodiversity, University of Münster, Münster 48149, Germany

[3]Department of Biology, University of Oxford, Oxford, UK

[4]Pandemic Sciences Institute, University of Oxford, UK

[5]Department of Microbiology, Immunology and Transplantation, Rega Institute, Laboratory for Clinical and Epidemiological Virology, KU Leuven, Leuven, Belgium

[6]Center for Interdisciplinary Research in Biology, CNRS, Collège de France, PSL Research University, Paris 75005, France

[7]Department of Biology, Georgetown University, Washington, District of Columbia, USA

[8]Department of Molecular Medicine, University of Padova, 35121 Padova, Italy

*Corresponding author. Email : chiara.poletto@unipd.it

[†]These authors contributed equally to this work

This PDF file includes :

- Supporting text
- Supplementary Fig. 1 to 5
- Supplementary Tables 1 to 5
- SI References

Supporting information Text

Arrival times, number of passengers and sequencing coverage

Assuming exponential increase of cases in the origin country, the time of first arrival with passenger flow p is Gumbel distributed with mean proportional to $log\ (p)$ [1]. Assuming collection coverage s on top, the mean arrival time based on first collection scales as $log\ (p)\ + log\ (s)$ and "iso-arrival-time" lines are as $log\ (p)\ + log\ (s)\ =\ k$ (anti diagonals). We report such lines in Fig. 1B as reference.

International dissemination model: Sensitivity analysis

In the sensitivity analysis we tested the following assumptions:

- Mean incubation period equal to 4 days
- Mean incubation period equal to 6 days
- Delays computed for each countries after the alert averaged over a sliding time window of 7 days
- Percentage of case detection outside the UK, K_c= 25%
- Flights from all England airports with a catchment population of 56 millions inhabitants
- No changepoint for the Alpha incidence exponential growth in the UK
- Two changepoints for the Alpha incidence exponential growth in the UK an 5 Nov 2020 and 2 Dec 2020

Supplementary Table 5 shows the results of the sensitivity analysis. For the baseline scenario and the sensitivity models tested we provide best estimates and some model predictions chosen as reference. Varying the parameters had little impact on the parameters estimated in the model. The number of countries with introduction before 31 Dec 2020 increased in the

following cases: delays from collection to submission for Alpha computed for each country aggregated over 7 days, 25% detection of imported cases, no change of slope.

Local dynamics in the USA at a finer spatial scale

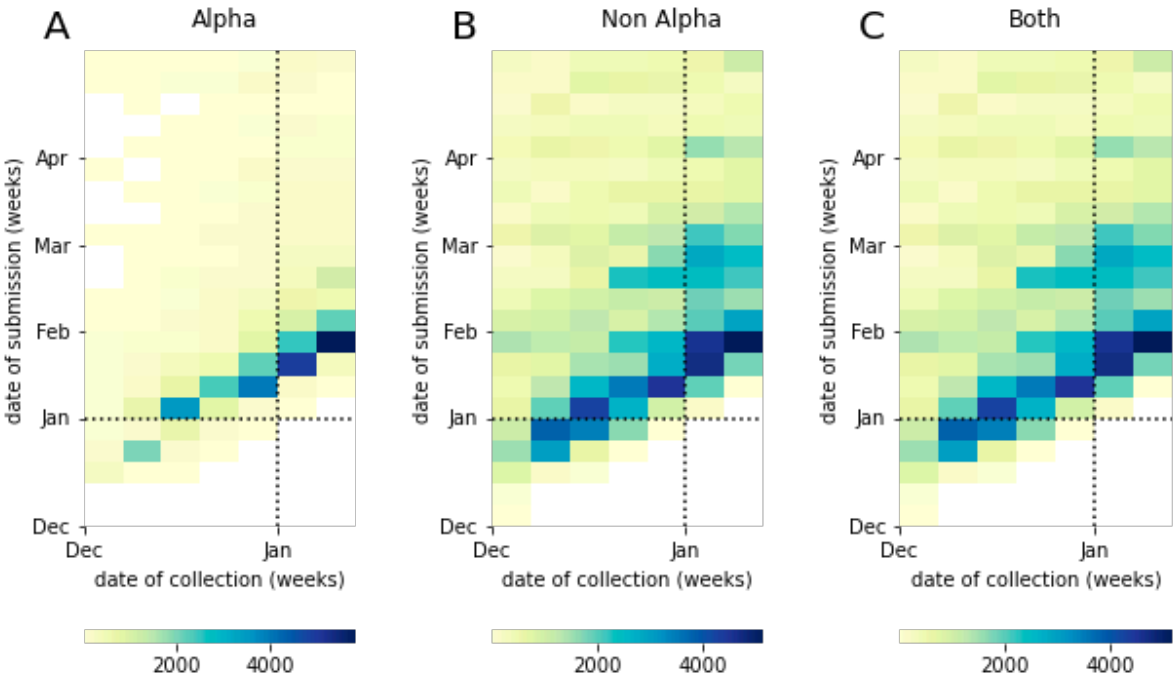
The analysis of Alpha local spread for the USA shows that this country is out of trend, with a high number of predicted Alpha cases compared with model estimates. Here, we carry out the comparison for two individual states, California and Florida, and for New York City - see sources of data reported in Supplementary Table 2. These locations were the port of entry of Alpha into the US, with early reported Alpha cases linked directly to the UK [2,3].

As for the USA as a whole, the autochthonous model A was fed with importation fluxes estimated from the international dissemination model and we compared the model-predicted number of Alpha cases with the empirical estimates. We present these results on Supplementary Fig. 4.

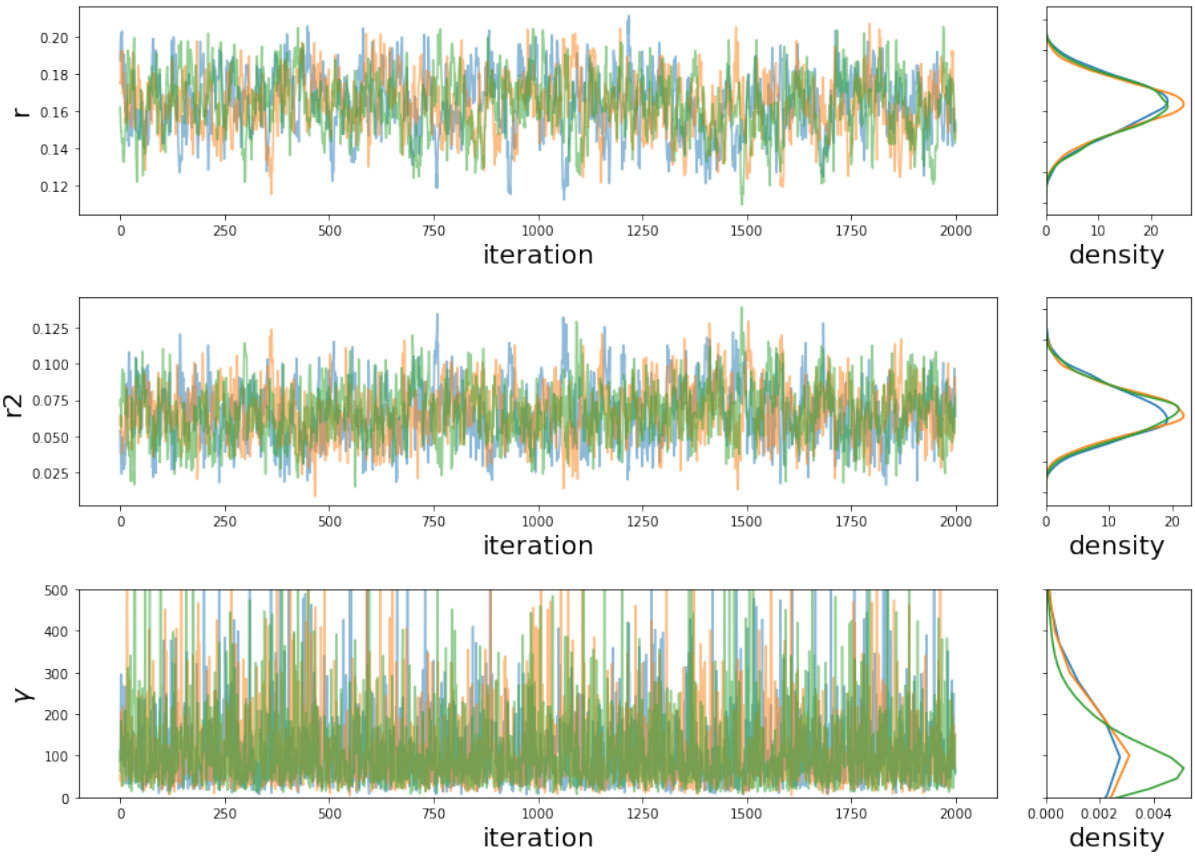
For California and New York City, the comparison between model and empirical estimates follows a trend similar to European countries. Florida registered a high proportion of Alpha cases [2]. Such a high level of Alpha circulation can be compatible with model predictions in a scenario of early Alpha introduction, i.e. introduction dates close to the lower bound of the range predicted by the model.

Median seeding time

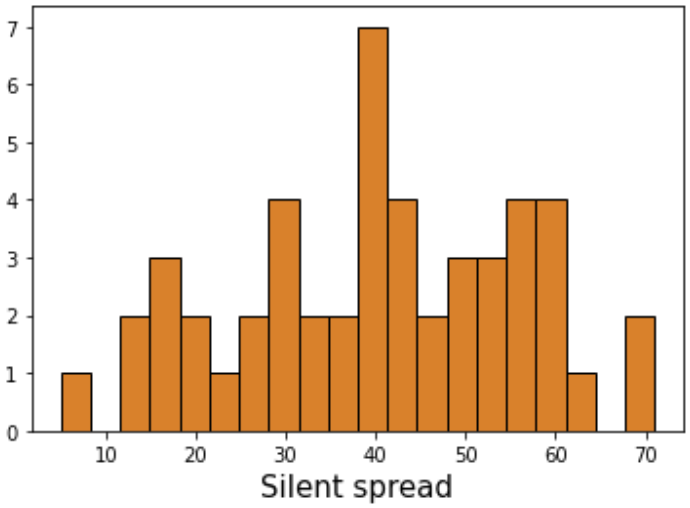
In the reference case in which traveling fluxes are constant in time and R_t is the same in the destination country as in the UK, the reference median date of seeding would fall halfway between the date of emergence and 31 Dec 2020 (see Methods). The median seeding dates of active chains at the end of 2020 departed from this assumed scenario. In Fig. 4D, we found that there was a negative correlation between the overall reproduction ratio over the period and the difference between computed median seeding date and reference median seeding date (spearman correlation -0.81 with p-value = 0.049), implying that lower transmissibility overall led to less success in early introductions. In Supplementary Fig. 5 we show that there was no significant correlation between the international traffic drop and the difference of the median seeding date with the reference date.



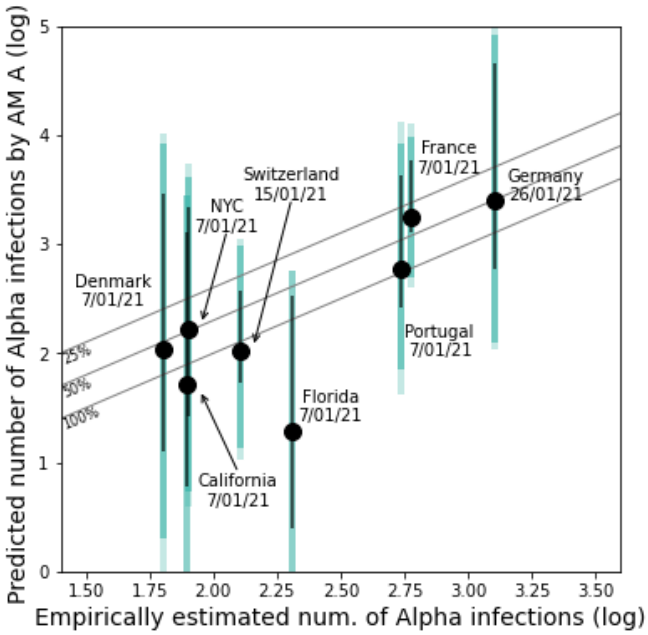
Supplementary Fig 1. Occurrences of delays between collection and submission in time. (**A**) Alpha variant.(**B**) Non Alpha variants. (**C**) Both.



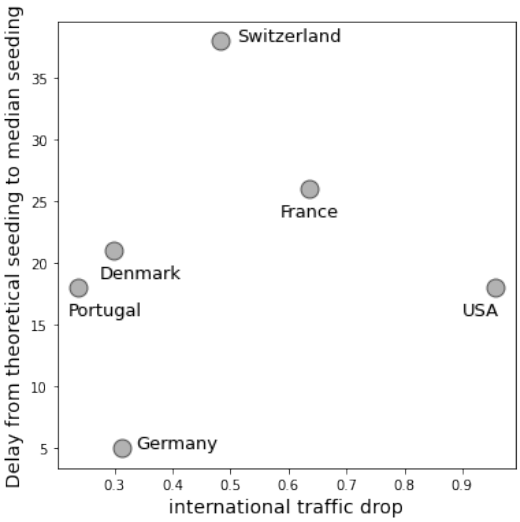
Supplementary Fig. 2. MCMC convergence plot. The fitted model has 2 exponential growth rates (r_1,r_2) with a changepoint on November 5th,2020. Three independent chains (red, green,blue) were run for 100000 iterations, with 50000 discarded as burn-in. Posterior samples were thinned 1 in 25.



Supplementary Fig. 3 Distribution of silent spread in days (n=69). Silent spread is computed as in Fig. 3D.



Supplementary Fig. 4. Model vs. empirical cases of Alpha as in Fig. 4 of the main paper. Here, the USA is replaced by California, New York City (NYC) and Florida to address spatial heterogeneity inside the USA. The empirical estimates of Alpha cases are computed by multiplying the Alpha frequency from virological investigations by the reported COVID-19 incidence at the same date - the date is indicated in the plot. Model estimates are obtained with the autochthonous model A (AM A in the plot). Gray lines show ratios of 100%, 50% and 25% between observed and predicted infections attributable to reporting. Black error bars indicate the prediction interval over 500 stochastic simulations obtained with the median volume of Alpha introduction, output of the international dissemination model assuming a 7-day delay between case and infection. Dark colored bars account for the variability in the output of the autochthonous model accounting for the upper and the lower limit of the prediction interval of the Alpha introductions as given by the international dissemination model. Light colored bars account for variability in the delay from infection to case reporting (ranging from 4 days to 10 days).



Supplementary Fig. 5. Difference between the median seeding date predicted by the autochthonous model A with the same quantity when R_t is the same in all countries and traveling fluxes do not change in time, plotted against the international traffic drop. The international traffic drop is computed as the international traffic in Nov 2020 divided by the average of international traffic between Sep 2020 and Oct 2020. The Spearman correlation coefficient does not show a correlation between the two quantities (coefficient = 0.23, p-value = 0.66, t-distribution with n-2 dof).

Supplementary tables

Country	Date of cases surveyed for Alpha	Daily number cases	Frequency Alpha	Computed number of Alpha infections	Source
France	7 Jan 2021	18,004	0.033	594	[4]
Portugal	4-10 Jan 2021	8,062 (7/01/2021)	0.068*	548	[5]
Germany	23-29 Jan 2021	12,370 (26/01/2021)	0.103	1,274	[6]
Denmark	4-10 Jan 2021	1,825 (7/01/2021)	0.035	64	[7]
Switzerland	15 Jan 2021	2,204	0.058	128	[8]
USA	7 Jan 2021	248,566	0.0048	1,193	[2]

Supplementary Table 1. Summary of data used for the validation. France epidemiological report last accessed 26/05/2023. Germany report last accessed 26/05/2023. Denmark website last accessed 26/05/2023 (2 Mar 2021 version).

* At week 01 of 2021, 7.38% of cases were suspicions of Alpha, 92% of which are true Alpha.

Country	Date of cases surveyed for Alpha	Daily number cases	Frequency Alpha	Computed number of Alpha infections	Source
California	7 Jan 2021	43314	0.0018	78	[2]
Florida	7 Jan 2021	15939	0.0128	204	[2]
NYC	7 Jan 2021	5808	0.013738	80	[9]

Supplementary Table 2. Summary of data used for the local spread analysis for the locations within the US explored in Supplementary Fig. 4.

Parameter	Description	Baseline value
K_{UK}	Fraction of sampled Covid cases.	0.25
K_c	Fraction of sampled imported Covid cases.	0.5
e	Incubation period.	5 days
N	Population in the catchment area of London airports	36M
T_0	Beginning of the risk window for VOC emergence in the UK	15 Aug 2020
$Tcp2$	Date of change of the exponential transmission growth in the UK.	5 Nov 2020

Supplementary Table 3. Summary of the parameters values assumed in the international dissemination model.

Parameter	Description	Prior distribution
r_1	Exponential growth rate in UK up to 5 Nov 2020	Exp(0.1)
r_2	Exponential growth rate in UK after 5 Nov 2020	N(0,1)
γ	Ratio of sampling among traveling vs. non-traveling cases after 18 Dec 2020	Exp(0.01)

Supplementary Table 4. Summary of the estimated parameters and their prior distribution.

Scenario	r_1	r_2	γ	Predicted time of emergence in the UK	Median date of first introduction for France, and Denmark	#countries with introduction before 31 Dec 2020
Baseline	0.17 [0.14;0.20]	0.055 [0.02;0.097]	51.67 [12.43;310.11]	08/09 [21/08;19/09]	17/10 [18/09-03/11] - 05/11 [08/10-06/12]	65 [52-73]
Incubation = 4 days	0.17 [0.14;0.20]	0.056 [0.024;0.096]	51.51 [11.72;307.19]	09/09 [21/08;18/09]	17/10 [20/09-03/11] - 05/11 [09/10-07/12]	65 [52-73]
Incubation = 6 days	0.17 [0.14;0.20]	0.056 [0.024;0.097]	52.36 [11.89;301.70]	06/09 [19/08;16/09]	17/10 [19/09-03/11] - 05/11 [09/10-06/12]	65 [52-73]
Delays computed by country, aggregated on 7 days After 18 Dec 2020	0.16 [0.11;0.19]	0.086 [0.043;0.13]	36.26 [7.60;186.32]	02/09 [17/08;18/09]	17/10 [20/09-09/11] - 09/11 [09/10-07/12]	69 [60-73]
25% detection of imported cases	0.18 [0.15;0.21]	0.058 [0.025;0.1]	51.20 [11.53;300.63]	06/09 [19/08;18/09]	15/10 [19/09-31/10] - 31/10 [07/10-30/11]	70 [61-73]
Air travel : flight from England	0.17 [0.14;0.20]	0.056 [0.025;0.097]	73.13 [15.46;415.13]	02/09 [17/08;18/09]	19/10 [21/09-06/11] - 07/11 [10/10-08/12]	64 [47-72]
No change of slope	0.12 [0.11;0.13]		22.46 [6.75;94.21]	30/08 [16/08;12/09]	23/10 [14/09-13/11] - 24/11 [12/10-13/12]	69 [59-73]
2 changes of slope : 11/5 and 12/2	0.16 [0.12;0.20]	r2 : 0.07 [0.026,0.15] r3 : 0.023 [0.022,0.48]	61.32 [13.19;332.14]	02/09 [17/08;17/09]	18/10 [18/09-08/10] - 08/11 [09/10-07/12]	65 [51-73]

Supplementary Table 5. Sensitivity analysis of the international dissemination model.

References

1. Gautreau, A., Barrat, A. & Barthélemy, M. Global disease spread: Statistics and estimation of arrival times. *Journal of Theoretical Biology* **251**, 509–522 (2008).
2. Washington, N. L. *et al.* Emergence and rapid transmission of SARS-CoV-2 B.1.1.7 in the United States. *Cell* **184**, 2587-2594.e7 (2021).
3. Alpert, T. *et al.* Early introductions and transmission of SARS-CoV-2 variant B.1.1.7 in the United States. *Cell* **184**, 2595-2604.e13 (2021).
4. SPF. COVID-19 : point épidémiologique du 28 janvier 2021. https://www.santepubliquefrance.fr/maladies-et-traumatismes/maladies-et-infections-respiratoires/infection-a-coronavirus/documents/bulletin-national/covid-19-point-epidemiologique-du-28-janvier-2021.
5. Borges, V. *et al.* Tracking SARS-CoV-2 lineage B.1.1.7 dissemination: insights from nationwide spike gene target failure (SGTF) and spike gene late detection (SGTL) data, Portugal, week 49 2020 to week 3 2021. *Eurosurveillance* **26**, 2100131 (2021).
6. RKI - Coronavirus SARS-CoV-2 - Berichte zu Virusvarianten von SARS-CoV-2 in Deutschland. https://www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/DESH/Berichte-VOC-tab.html.
7. Danish Covid-19 Genome Consortium. https://covid19genomics.dk/statistics.
8. Chen, C. *et al.* Quantification of the spread of SARS-CoV-2 variant B.1.1.7 in Switzerland. *Epidemics* **37**, 100480 (2021).
9. coronavirus-data/variants at master · nychealth/coronavirus-data. *GitHub* https://github.com/nychealth/coronavirus-data/tree/master/variants.

3.3 Conclusion.

In this chapter, using metadata from sequences of the Alpha variant submitted to GISAID, we explored the interplay between sequencing coverage, air travel volumes on the duration of silent circulation. The sequencing coverage, as well as the delay between the collection and submission of a genomic sequence, varied by several orders of magnitude between countries. The volume of air traffic between the source of the variant (UK) and other countries showed similar heterogeneity. A broader analysis, considering sequences submitted to GISAID between December 1, 2019, and October 31, 2021, confirmed this disparity [2]. The authors of this study identified two main sources of sequencing inequalities. First, they showed that the proportion of sequenced cases was strongly correlated with the economic development level of the country. Low-income countries were more likely to share fewer sequences than high-income countries due to the lack of infrastructure. However, heterogeneity was also observed across high-income countries. These differences were mainly due to different testing policies. The latter point emphasised the importance of a coordinated sequencing policy among countries. Throughout the pandemic, international agencies such as ECDC [144, 145] and WHO [146] had provided guidelines to standardise sequencing practices and improve data accessibility. These reports particularly emphasised the need to provide high-quality data with standardised metadata. Once collected, sequences should be publicly shared on platforms easily accessible to stakeholders worldwide.

Due to these heterogeneities, surveillance data were biased and incoherent. By integrating all factors at play we estimated that the Alpha variant had circulated in most countries undetected for a period of up to two months. More than 60 countries had likely experienced an introduction of the Alpha variant before the end of December. Overall, our analysis showed that the virus propagation was largely driven by passenger flows from the source country, which confirmed previous studies [38, 53, 104]. At the same time, however, the intensity of sequencing efforts mainly determined the duration of silent circulation.

Our study provides a model that can integrate data from different sources to reconstruct the epidemic unfolding. The different components involved in viral dissemination and detection have been explicitly modelled: growth in the UK, exportation to another country, detection, and delay in sample collection to submission. We could expand the analysis and test different hypotheses regarding sequencing and air travel would be possible using counterfactual scenarios. We could determine the minimum sequencing intensity required to detect a variant within a specified timeframe since its emergence or investigate whether an earlier reduction in air travel could have delayed the exportation of the variant. This framework could also be used to study the spread and the detection of other variants, or future viruses.

While our study is retrospective, a similar framework could be used to analyse the epidemic unfolding in real-time and forecast the propagation of a new variant under certain conditions. Firstly, our framework needs real-time estimation of sequencing coverage across different countries as input data. Sequencing coverage computed from the preceding months before such a study could be an acceptable proxy. Secondly, it is worth noting that our model assumes variant exportation occurs directly from the source country to other countries. While other studies have validated this assumption for the case of Alpha for the period under study, it implies that the model may only apply to viruses/variants with a single source and during the initial months of propagation. For instance, as discussed in the first chapter of this thesis, such an assumption couldn't be applied to the Delta or Omicron variant [38]. We could complexify the model to allow exportation from a non-primary source country to relax this

assumption.

Chapter 4

Agent-based modelling of reactive vaccination of workplaces and schools against COVID-19.

In this Chapter, I present the work published in [6]. In this study, we used an agent-based model to assess the effectiveness of reactive vaccination to mitigate COVID-19 spread and control the emergence of a new variant. Based on historical examples, I review the different reactive vaccination strategies and the different epidemiological contexts in which they were used. Then, I present the details of the agent-based model used in the study, and I will present the article.

4.1 Introduction.

4.1.1 Target vaccination.

Historical examples have demonstrated the effectiveness of vaccination in reducing cases of infectious diseases [147] and even eliminating (polio or diphteria) or eradicating (smallpox) them in some parts of the world [148]. The basic theory [149] estimates the proportion of the population p to be vaccinated based on the basic reproduction number $p = \frac{1}{VE}(1 - \frac{1}{R0})$, where VE is the vaccine efficacy, and R_0 is the basic reproduction number. However, this theory relies on simplifying assumptions, such as homogeneous mixing. The proportion p might change if contacts or susceptibility are heterogeneous.

Choosing the right vaccination strategy was critical during the first semester of 2021. Vaccine distribution, while limited by the number of available doses, had to reduce the incidence of COVID-19 in the context of pandemic fatigue [11], reducing the impact of social distancing measures. In the context of the Delta wave in summer 2021, vaccination was employed as a strategy to limit the rise of cases. In particular, a more targeted or reactive vaccination aimed at vaccinating those at the highest risk of being infected was used occasionally to maximise the impact of the initial vaccine doses distributed. In May 2021, the UK government chose to increase testing and vaccine distribution in areas where an increase in cases related to the spread of the Delta variant was detected [150]. On a smaller scale, Cologne [151] massively vaccinated neighbourhoods with detected hotspots, regardless of individuals' ages. In France, targeted vaccination was implemented in Bordeaux, Strasbourg, and Brest [7] to contain the emergence of the Delta variant. In Strasbourg, vaccine doses were distributed to students at an art school after a cluster of 4 cases of the Delta variant was detected [152].

Several successful reactive vaccination strategies have been implemented in the past. Modelling studies assessed the effectiveness of vaccination of close contacts for smallpox [132] and vaccination of children in urban areas where an outbreak was detected for measles [153]. Ring vaccination was implemented in the context of the Ebola virus endemic in West Africa in 2014-2015, with contacts and contacts of contacts of detected cases being vaccinated. Randomised controlled trials assessed this reactive intervention's effectiveness [153, 154].

By targeting the contacts of detected cases, ring vaccination was expected to interrupt transmission chains early, limiting the spread of the disease or even potentially leading to its eradication [155]. Furthermore, ring vaccination was supposed to decrease the vaccine doses needed compared with a traditional mass vaccination campaign, which can be useful early in an epidemic when the total supply of vaccine doses is unavailable [155]. However, this strategy may not be suitable for all cases. The eradication of smallpox using this strategy was possible because the disease was easily identifiable with recognisable symptoms, without a significant latent or asymptomatic infectious phase [156]. Additionally, the vaccine was cheap, readily available, and stable at high temperatures [156], with high effectiveness at preventing infection and transmission. The vaccine was also fast to be effective after injection compared to the long incubation period. Ring vaccination would be ineffective for diseases with a significant pre-symptomatic phase during which an undetected individual could cause secondary cases. It also requires an effective system for case detection and contact tracing, adherence of contacts to getting vaccinated, and a sufficiently agile logistics system for rapid deployment. A modelling study of the impact of ring vaccination in the context of an Ebola epidemic in Sierra Leone [157], using an agent-based model, highlighted critical factors for the success of this strategy. In all scenarios, the effectiveness of reactive vaccination decreased as transmissibility, human mobility, or the time between infection and hospital admission increased. The overall rate of vaccine coverage had little impact.

Another type of reactive vaccination involves vaccinating all individuals in a location (e.g., workplace, school) identified as a hotspot of the epidemic. This strategy has been used several times to address resurgences of measles cases in past decades. Between October 2012 and September 2013, multiple measles outbreaks were identified in the young population of the Greater Manchester area [158]. The UK authorities targeted secondary schools where cases had been detected to contain measles spread. These local vaccinations were accompanied by information campaigns to increase the adherence of unimmunised families to the vaccination program. A similar strategy was used in Edinburgh among students when outbreaks occurred in secondary schools and universities in the city. In the case of the 2003-2004 measles epidemic in Niamey in Niger, reactive vaccination was found to be the optimal strategy in case of low initial vaccine coverage. In the context of the COVID-19 pandemic during the first semester of 2021, we conducted a study to explore the feasibility and impact of implementing a reactive vaccination strategy in France. Among the various possible forms of reactive vaccinations, we examined a strategy targeting schools and workplaces where cases were detected. Ring vaccination was unlikely to be useful for COVID-19 for several reasons. First, a significant proportion of cases were asymptomatic or have mild symptoms. A meta-analysis from 2020 demonstrated that 24.51% of transmissions were asymptomatic [159]. Even among symptomatic cases, the majority went undetected by the surveillance system. For example, in France, it was shown that nine out of ten symptomatic cases went undetected during the first wave in spring 2020 [9]. Furthermore, even for symptomatic cases, there is a pre-symptomatic phase

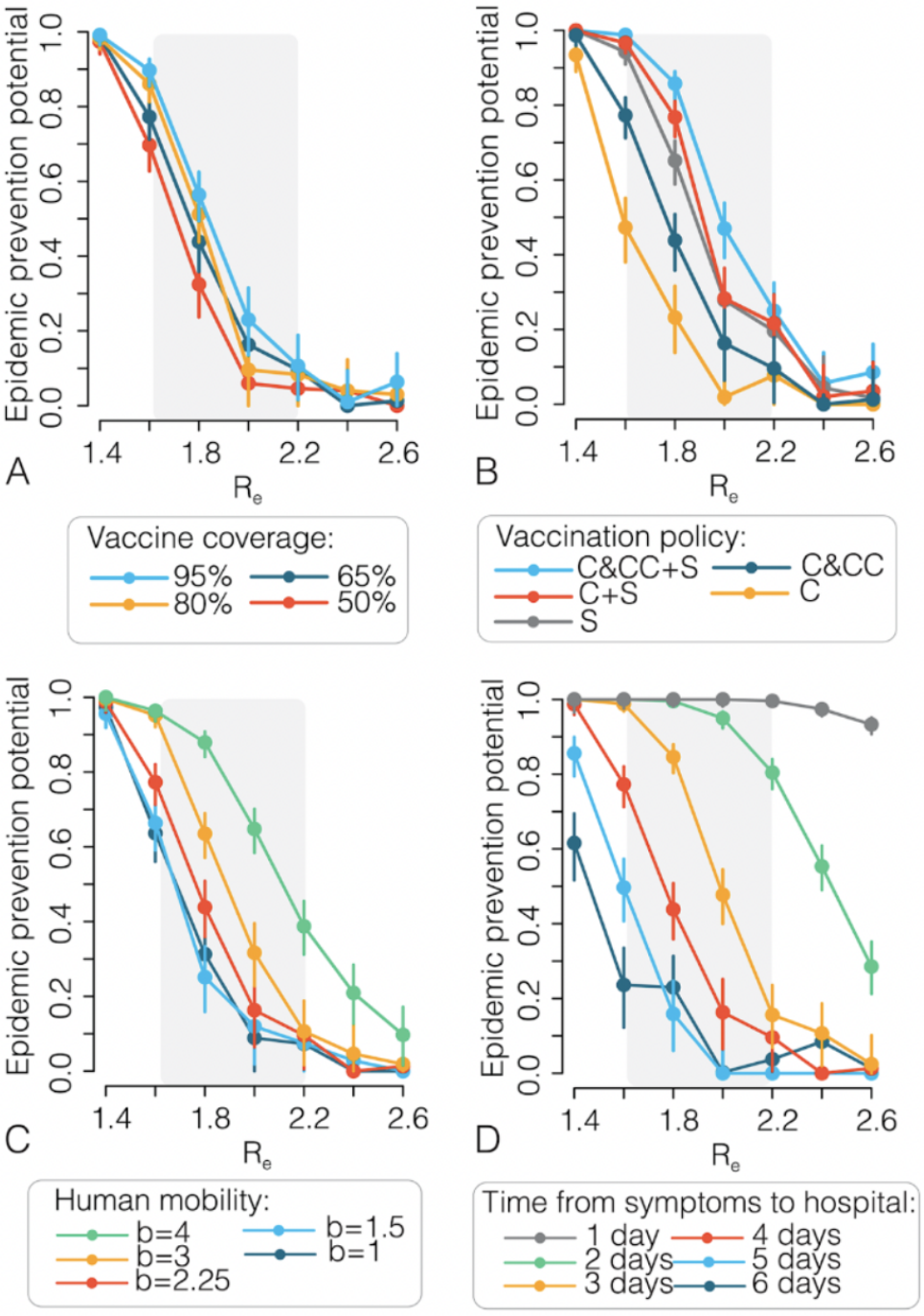


FIGURE 4.1: From [157]. Epidemic prevention potential of ring vaccination to contain Ebola epidemic in Sierra Leone. Various parameters are tested. (A) Vaccine coverage. (B) Vaccination policy: C (Contact of index), CC (contact of contact of index), S (Geographical rings of 2 km). (C) Human mobility Spatial transmission is proportional to a power law kernel 1/(1 + db) where d is the geographical distance and b regulates the decrease of transmission with distance. (D) Time from symptoms to hospital.

during which individuals are infectious and can generate secondary cases before developing symptoms. The infectious period has been estimated to start days before the onset of symptoms [160]. Additionally, even after an individual develops symptoms, it takes several days for them to be detected by the health insurance system -on average, 2.6 days in France in the first half of 2021 [161]. Considering that the vaccine

takes around two weeks [71] to become effective, a first-degree reactive ring vaccination would not be effective in reducing cases. It could have been possible to study the effect of second-degree ring vaccination, where contacts of contacts are vaccinated. However, this strategy would have raised many logistic issues and would have been challenging to implement. Thus, it would have had a limited public health interest. Vaccinating schools or workplaces of index cases allows for the easy targeting of the contact network of cases, as most of an individual's contacts were found to occur in these settings [116]. Our study investigated the impact of such a vaccination strategy in the context of the pandemic during the first half of 2021. At this time, several restrictions were in place (curfew, social distancing measures, teleworking), and vaccine distribution was limited to the elderly population. On the other hand, many countries had already implemented TTI systems, allowing for detection of cases, tracing of contact and isolation. The implementation of reactive vaccination could leverage these existing infrastructures. Classical compartmental models have already been used to address public health questions about vaccination [76, 162, 163, 164]. However, an agent-based model would be more appropriate to study reactive vaccination as it requires understanding the interplay between different time scales (detection of cases, isolation, contact tracing, time necessary for the vaccine to become effective) at the individual level. Our agent-based model was an extension of a previous one used to assess the impact of contact tracing apps [59]. The stochastic nature of the model captured the inherent randomness in the spread of the epidemic and incorporated uncertainty in model parameters (network structure, seeding, and variations in contacts). Our work aimed to investigate the impact of various epidemiological factors on the effectiveness of reactive vaccination strategy, including population immunity, virus transmissibility, initial vaccine coverage, vaccine efficacy, and public adherence to vaccination. We quantified the reduction in symptomatic cases and the daily number of vaccine doses administered. We also examined the possibility of controlling an outbreak at the start after detecting a local cluster to assess the strategy's effectiveness against a new variant. Finally, if a reactive vaccination strategy was implemented, it would likely overlap with a mass (e.g. non-targeted) vaccination. Then, to consider a realistic scenario, we also studied a mixed situation where both strategies are employed simultaneously to understand their interactions.

4.2 Details about the model.

4.2.1 Modelling population.

We extended the model described in [58] to model the spread of COVID-19 in a population under social distancing restrictions. The synthetic population was stochastically built using demographic data from Institut National de la Statistique et des Etudes Economiques (INSEE). We reproduced statistics on age distribution, household number and type, and school and workplace number and size. We used Metz, a middle-sized city in the Grand Est region, as a case study. Different types of households were considered, depending on whether they were composed of a couple or a single individual, with or without children. Each household type had a certain probability of being composed of individuals of a certain age. At the start of a simulation, a set of households is randomly generated, with their composition and type based on the distributions observed in INSEE data, until the total population size of Metz (117,492 inhabitants) is reached [59]. Once the population is built, each individual is randomly assigned to a school, university, or workplace based on their age. We modelled daily contacts between individuals with a time-varying contact network with five layers,

each representing a location where contacts occur: school, workplace, home, community, and public transport. For each simulation, we built an acquaintance network formed by the combination of Erdos-Renyi networks in each location (e.g., each school and workplace). The average degree $\chi_i(s,n)$ is a random variable that depends on the type of location s and its size n. We sample its value from a gamma distribution with an average $\bar{\chi}_i(s,n)$ and a variation coefficient CV. We assume that CV is the same for all settings. We parametrise $\bar{\chi}_i(s,n)$ so that for small sizes, all individuals come into contact, whereas as the size increases, the number of contacts saturates. We denote with w_s this maximum of contacts.

$$\bar{\chi}_i(s,n) = \frac{(n-1).w_s}{w_s + (n-1)}$$

Once the network is constructed, each link is assigned a daily activation rate x, sampled from a cumulative distribution function $F_s(x)$. This function is a sigmoid with parameters A_s and B_s . Parameters A_s,B_s,w_s and CV are fixed to reproduce the statistics of daily contacts reported in [116]. Contacts are then sampled each day according to the activation rate.

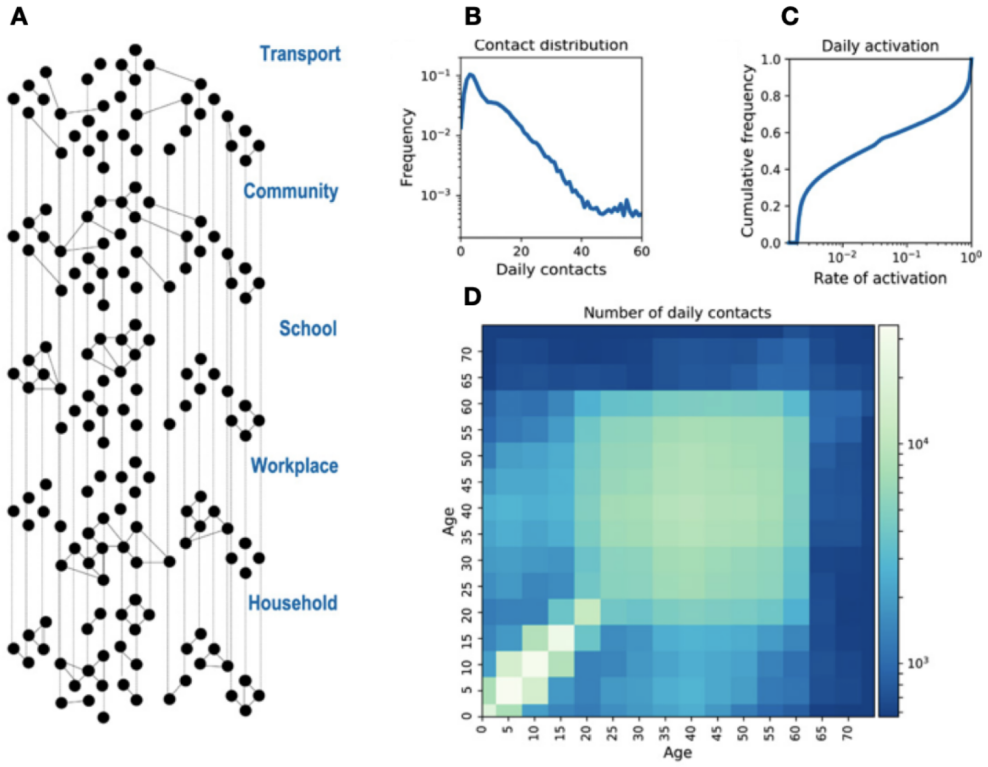


FIGURE 4.2: (from [59]): Synthetic population used in the model. (A) Construction of the multilayer contact network. Contacts among individuals were represented as a multilayer dynamical network, where each layer includes contacts occurring in a specific setting. (B) Distribution of the number of daily contacts in the model. (C) Cumulative distribution of the activation rate associated with the contacts in the model was calibrated to be consistent with the information from INSEE. (D) Age contact matrix computed from the contact network model.

4.2.2 Modelling social distancing.

To analyse the impact of social distancing measures on the vaccination strategy, we modelled contact tracing and isolation measures for detected cases and their contacts. Once infected, an individual is detected with a probability $p_{d,c}$ if he is clinical or with a probability $p_{d,sc}$ if he is subclinical and then isolated with a rate r_d. Household contacts of index cases are isolated with a probability $p_{ct,HH}$. As soon as an index case is detected, a proportion of contacts outside the household is traced and isolated with a mean delay r_{ct}^{-1}. Frequent contacts are distinguished from the ones occurring more rarely. To do so, we define a threshold f_a such that cases frequently in contact with the index case ($f \geq f_a$) are detected with a probability $p_{ct,A}$ while others are detected with probability $p_{ct,sp}$. Only contacts that occurred D days before the detection of the index case are traced. The duration of isolation depends on the infectious status of the individual, as legislation could have differed between actual positive cases and non-positive contacts. If infected, the individual stays isolated for a period d_I. Otherwise (Susceptible or Recovered), he stays isolated for a period d_{NI}. Asymptomatic contacts have a probability p_{drop} of dropping out of isolation each day.

The model also accounts for teleworking and social distancing. At the beginning of each simulation, a portion of contacts is reduced in workplaces (to represent teleworking) and the community (for social distancing measures). These contact reductions remain constant throughout the simulated period - around two months. In the baseline scenario analysed in the paper, we set the percentage of contacts removed in workplaces at 10% and 5% in the community to match the values observed in France during the autumn of 2021 and reported by Google Mobility data [165].

4.2.3 Epidemiological model.

The transmission of COVID-19 and individual infection history were described by assigning each individual a status with susceptibility and severity depending on age and vaccination status. An individual who has not yet encountered the disease is Susceptible (S). Once infected, he remains exposed (E) for a random latency period ϵ^{-1} following en Erlang distribution with an average of 3.7 days [166], during which he is neither infectious nor symptomatic. The individual then become infectious and can infect other susceptible individuals. To model the variation in symptom intensity, the infectious individual can be subclinical (corresponding to asymptomatic or mildly symptomatic cases) or clinical (with moderate to severe symptoms). The probability p_{sc}^{A} [167] of being subclinical is dependent on age A. Before entering the I_c or I_{sc} compartments, the infectious individual go through a pre-symptomatic phase with an average duration μ_p^{-1} of 2.1 days [166]. An infectious individual recovers and enters the Recovered (R) compartment with an average duration μ^{-1} of 7 days. A recovered individual is no longer infectious and can not be infected again. We did not modelled waning immunity as we simulated an epidemic for a maximum of two months. When an individual is susceptible, he can be infected by another infectious individual with a probability per contact and per unit of time given by the product $\beta.\beta_I.\omega_S.\sigma_A$. The term β_I represents the reduction in transmissibility of pre-symptomatic individuals (β_I=0.51) compared to symptomatic individuals (β_I=1) [114]. This value is independent of whether the individual is clinical or subclinical. The term ω_s allows for variability in transmissibility based on location ($\omega_s = 1$ for households, $\omega_s = 0.3$ for community contacts and $\omega_s = 0.5$ in other places) [59]. Lastly, the term σ_A reflects the variability in susceptibility of the infected individual based on their age [59]. The value of β was set for each scenario so that the effective reproduction number R_{eff} at

the onset of the simulation is at a value of 1.6, corresponding to measurements for the Delta variant [168]. We computed the R_{eff} as the average number of secondary cases generated by an index case in a scenario without vaccination.

We assumed a "leaky" vaccination to model the effect of vaccination. All vaccinated people were partially protected against infection, transmission, and clinical symptoms. We added compartments for vaccinated individuals to the SEIR-like model, allowing for protection against infection and the development of moderate to severe symptoms. Once a susceptible individual is vaccinated, he enters the $S^{V,0}$ compartment where he remains for an average time τ_0, representing the delay in the first dose's effectiveness. The individual then moves to the $S^{V,1}$ compartment, corresponding to full single-dose vaccine efficacy. After an average time τ_1 (the time needed to get the second dose and for this to become effective), the individual entered the $S^{V,2}$ compartment where the two-dose vaccination was fully effective. For $S^{V,2}$ compartment, we denote with $r_{s,2}$ the reduction in susceptibility to infection and with $r_{c,2}$ the reduction in the probability of developing clinical symptoms. Thus, we have:

$$r_{s,2} = \frac{P(S^{V2} \longrightarrow E^{V2})}{P(S \longrightarrow E)} = 1 - VE_{s,2},$$

and

$$r_{c,2} = \frac{P(E^{V2} \longrightarrow I^{V2}_{p,c})}{P(E \longrightarrow I_{p,c})}$$

with

$$1 - VE_{sp,2} = \frac{P(I_c|V)}{P(I_c|\bar{V})} = \frac{P(I_c|V,E).P(E|V)}{P(Ic|\bar{V},E).P(E|\bar{V})} = r_{c,2}.(1 - VE_{s,2}).$$

Here V designates a vaccinated individual and $\bar{V}$ a non-vaccinated one.
From the latter equation, we thus obtain,

$$r_{c,2} = \frac{1 - VE_{sp,2}}{1 - VE_{s,2}}$$

For S^{V1} compartment, we assume for simplicity a polarised effect, with a part of the vaccinated population behaving as fully vaccinated (Figure 4.3). We note r_{S1} the reduction in susceptibility and p_v the probability for an infected individual to go to compartment E^V. A computation similar to the computations above gives :

$$1 - VE_{sp,1} = (1 - VE_{s,1}).(p_v.r_{c,2} + (1 - p_v))$$

We assumed no reduction in transmissibility when a vaccinated person infects another, but we assume that the infectious period is reduced by 25%. We summarise the entire compartmental model in Figure 4.3.

We chose the vaccine effectiveness to match those observed for the BNT162b2 vaccine against the Delta variant. In the results section of the paper presented in this chapter, we detail these values and their corresponding sources.

4.2.4 Strategies and outcomes.

We compared reactive vaccination to other strategies. We assumed all these strategies were implemented when a given proportion of elderlies had already been

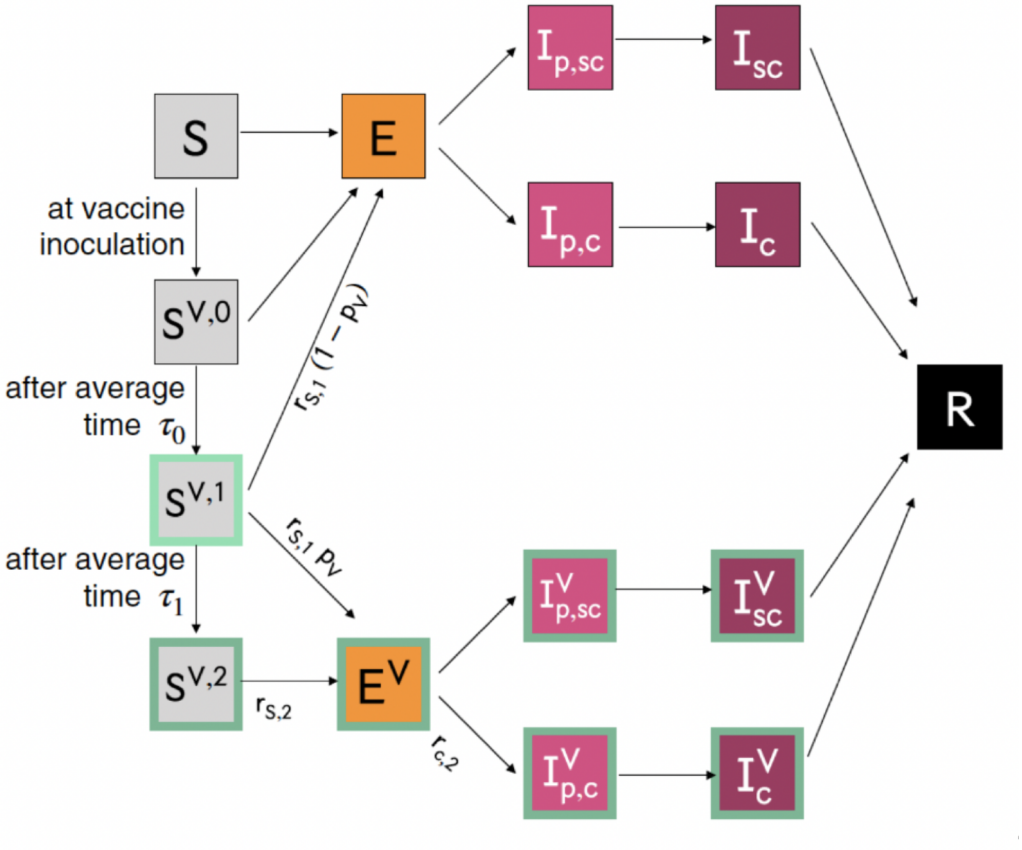


FIGURE 4.3: Compartmental model of COVID-19 transmission and vaccination.

vaccinated. In all scenarios, only individuals willing to get the vaccine were vaccinated. We parametrised the compliance to vaccination based on statistics [169]. We describe all strategies in the following.

Reactive vaccination: When a cluster of cases is detected in a workplace or a school, the entire setting is vaccinated at a rate r_v, as well as the household members of people being in that setting. We defined a cluster as a number of cases greater than or equal to n_{cl} over a period of T_{cl}. In some scenarios, we assumed the maximum number of doses administered per day (V_{daily}) and over the entire simulation duration (V_{tot}) was caped.

- *Baseline scenario*: No vaccination during the simulation.

- *Mass vaccination*: Every day, a set of individuals is randomly vaccinated. This strategy corresponds to a vaccination setup in vaccination centres, by GPs, or pharmacists.

- *School vaccination*: Every day, schools are randomly selected to vaccinate children attending the school, along with the household members of the children.

- *Workplace/University vaccination*: Every day, workplaces or universities are randomly selected to vaccinate their population and household members. School or workplace/university vaccinations are targeted vaccination but non-reactive strategies. By comparing their results to the reactive approach, we can explore the specific advantage of reactivity. *Combined strategy*: If reactive vaccination was implemented, it would likely be carried out in addition to a mass vaccination strategy, with individuals vaccinated daily at vaccination centres. Therefore, we also explore the possibility of a combination of mass vaccination and reactive vaccination. The number of detected

cases determined the number of doses for reactive vaccinations. We assumed a certain number of doses was distributed daily for non-reactive vaccinations.

At the beginning of each simulation, we set population immunity. We explored each strategy's ability to reduce the attack rate and the peak incidence compared to the scenario without vaccination. We also assessed the feasibility of reactive vaccination by measuring the doses distributed daily.

4.3 Article: Agent-based modelling of reactive vaccination of workplaces and schools against COVID-19.

nature COMMUNICATIONS

ARTICLE

https://doi.org/10.1038/s41467-022-29015-y OPEN

Agent-based modelling of reactive vaccination of workplaces and schools against COVID-19

Benjamin Faucher[1], Rania Assab[1], Jonathan Roux[2], Daniel Levy-Bruhl[3], Cécile Tran Kiem[4,5], Simon Cauchemez[4], Laura Zanetti[6], Vittoria Colizza[1,7], Pierre-Yves Boëlle[1] & Chiara Poletto[1✉]

With vaccination against COVID-19 stalled in some countries, increasing vaccine accessibility and distribution could help keep transmission under control. Here, we study the impact of reactive vaccination targeting schools and workplaces where cases are detected, with an agent-based model accounting for COVID-19 natural history, vaccine characteristics, demographics, behavioural changes and social distancing. In most scenarios, reactive vaccination leads to a higher reduction in cases compared with non-reactive strategies using the same number of doses. The reactive strategy could however be less effective than a moderate/high pace mass vaccination program if initial vaccination coverage is high or disease incidence is low, because few people would be vaccinated around each case. In case of flare-ups, reactive vaccination could better mitigate spread if it is implemented quickly, is supported by enhanced test-trace-isolate and triggers an increased vaccine uptake. These results provide key information to plan an adaptive vaccination rollout.

[1] Sorbonne Université, INSERM, Institut Pierre Louis d'Epidémiologie et de Santé Publique, Paris, France. [2] Univ Rennes, EHESP, CNRS, ARENES—UMR 6051, F-35000 Rennes, France. [3] Santé Publique France, Saint Maurice, France. [4] Mathematical Modelling of Infectious Diseases Unit, Institut Pasteur, Université de Paris, UMR2000, CNRS, Paris, France. [5] Collège Doctoral, Sorbonne Université, Paris, France. [6] Haute Autorité de Santé, Saint-Denis, France. [7] Tokyo Tech World Research Hub Initiative (WRHI), Tokyo Institute of Technology, Tokyo, Japan. ✉email: chiara.poletto@inserm.fr

Vaccination against SARS-CoV-2 has changed the course of the COVID-19 pandemic due to the high efficacy of available vaccines in preventing infection and severe disease. Yet, several months into the vaccination campaign, vaccine uptake remained below official targets in many Western countries due to logistical issues, vaccine accessibility and/or hesitancy. As of Fall 2021, less than 60% of the population in the United States and Europe was fully vaccinated[1]. With intense virus circulation still ongoing in many regions of the world due to the Delta variant and the threat posed by emerging variants, it is important to investigate whether vaccine use could improve with adaptive delivery. Indeed, offering vaccination to individuals who were exposed to the virus allows targeting those at higher risk of infection and, furthermore, might help overcome barriers to vaccination[2,3] since vaccine-hesitant people are more likely to accept vaccination when the perceived risk of infection is higher[4].

Redirecting vaccine supplies to geographic areas of highest incidence (or hotspot vaccination) is already part of the plans in some European countries and was implemented to combat the emergence of the Delta variant[2]. Other reactive vaccination schemes are possible, such as ring vaccination that targets contacts of confirmed cases or contacts of those contacts, or vaccination in workplaces or schools where cases have been detected. This could potentially improve the impact of vaccination by preventing transmission where it is active and even enable the efficient management of flare-ups. For smallpox or Ebola fever, ring vaccination has proved effective in rapidly curtailing outbreaks[5–8]. However, the experience of these past epidemics cannot be transposed directly to COVID-19 due to the many differences in the infection characteristics and epidemiological context. For example, COVID-19 cases are infectious a few days before symptom onset[9] but often detected a few days after. This gives time to infect their direct contacts and thwarts ring vaccination. Vaccinating an extended network of contacts, as could be done with the vaccination of whole workplaces or schools, would have a larger impact, especially if adopted in combination with strengthened protective measures to slow down transmissions, such as masks, physical distancing and contact tracing. This could be feasible in many countries, leveraging the established test-trace-isolate (TTI) system that enables prompt detection of clusters of cases to decide where vaccines should be deployed. Properly assessing the interest of reactive vaccination therefore requires a detailed examination of the interactions of vaccine characteristics, the pace of vaccination, COVID-19 natural history, case detection practises and overall changes in population behaviour.

We therefore extend an agent-based model that has been previously described[10] to quantify the impact of a reactive vaccination strategy targeting workplaces, universities and 12+ years old in schools where cases have been detected. We compare the impact of reactive vaccination with non-reactive vaccination targeting similar settings or with mass vaccination, and test these strategies alone and in combination. We explore differences in vaccine availability and logistical constraints, and assess the influence of the dynamic of the epidemic and different stages of the vaccination campaign.

Results

Mass vaccination, targeted and reactive vaccination strategies. We extended a previously described SARS-CoV-2 transmission model[10] to simulate vaccine administration alongside other interventions—i.e. contact tracing, teleworking and social restrictions. Following similar approaches[11–14], the model is stochastic and individual-based. It takes as input a synthetic population reproducing demographic and social-contact data, workplace sizes and school types (Fig. 1a) of a typical medium-sized French town (117,492 inhabitants). Contacts are described as a dynamic multilayer network[10] (Fig. 1b).

We assumed that the vaccine reduced susceptibility, quantified by the vaccine effectiveness VE_S, and symptomatic illness after infection, quantified by VE_{SP}[15] (Fig. 1c). We considered a vaccination strategy based on the Cominarty vaccine[16] which is very suitable for reactive vaccination given efficacy, only 3 weeks between the two doses and wide availability. We described the vaccine-induced protection with respect to the Delta variant—i.e. the dominant variant as of Fall 2021. Real-life estimates are heterogeneous, reflecting the complex interplay between the timing of Delta introduction in the population, the co-circulation of other variants, waning of immunity and differential impact by age. In the baseline scenario we considered vaccine effectiveness levels in the middle of the range of estimates provided in a systematic review[17]. We used a three-week interval between doses as in the vaccine trial[16]. For vaccine protection, we conservatively assumed that there was no protection in the 2 weeks after the first inoculation, followed by intermediate protection until 2 weeks after the second dose ($VE_{S,1} = 48\%$ and $VE_{SP,1} = 53\%$, see additional details in the Supplementary Table 2) and maximum protection afterwards ($VE_{S,2} = 70\%$ and $VE_{SP,2} = 73\%$), 5 weeks after the first dose. The maximum protection values are close to the estimates obtained in a meta-analysis for Delta, all vaccines combined[18]. Lower and higher vaccine effectiveness are also explored.

In the baseline scenario we parametrised the epidemiological context assuming that 32%[19,20] of the population was fully immune to the virus due to the previous infection. Initial incidence was moderate/high, i.e. ~160 clinical cases weekly per 100,000 inhabitants, and the reproductive ratio was $R = 1.6$, in the range of values estimated for the Delta wave of summer 2021[1,21]. We modelled the baseline TTI policy after the French situation, allowing 3.6 days on average from symptoms onset to detection and 2.8 average contacts detected and isolated per index case[22] (Fig. 1d). We assumed that 50% of clinical cases and 10% of subclinical cases were detected, leading to an overall detection rate of ~25%[20,23]. Social restrictions were modelled assuming 10% of individuals were doing teleworking and contacts in the community were reduced by 5% (see Methods).

We then modelled vaccination targeting all adults older than 12 years old with baseline vaccine uptake—set to 80% in the 12-65 years old and 90% in the over 65 years old[24]. We assumed that priority risk groups (e.g. elderlies) had already been vaccinated up to that level at the start[25–28]. We modelled three non-reactive vaccination strategies in the general population, where vaccination was carried out up to the maximum number of doses available each day at random (i) in the whole population (mass) or (ii) in schools sites (school locations, described below) or (iii) in workplaces/universities (workplaces/universities). In the school locations vaccination, we assumed vaccine sites were set up in schools to vaccinate pupils and their parents/siblings over the age of 12[29]. Then, we modelled a reactive vaccination strategy, where the detection of a case thanks to TTI triggered the vaccination of household members and those in the same workplace or school (Fig. 1d). In this scenario, a baseline delay of 2 days on average was assumed between the detection of the case and starting vaccination to account for logistical issues—i.e. ~5.6 days on average from the index case's symptoms onset. In the baseline scenario, we assumed vaccine uptake in the context of reactive vaccination to be the same as in non-reactive vaccination. The impact of each strategy was assessed by comparison with a reference scenario, where no vaccination campaign is conducted during the course of the simulation and vaccination coverage remains at its initial level.

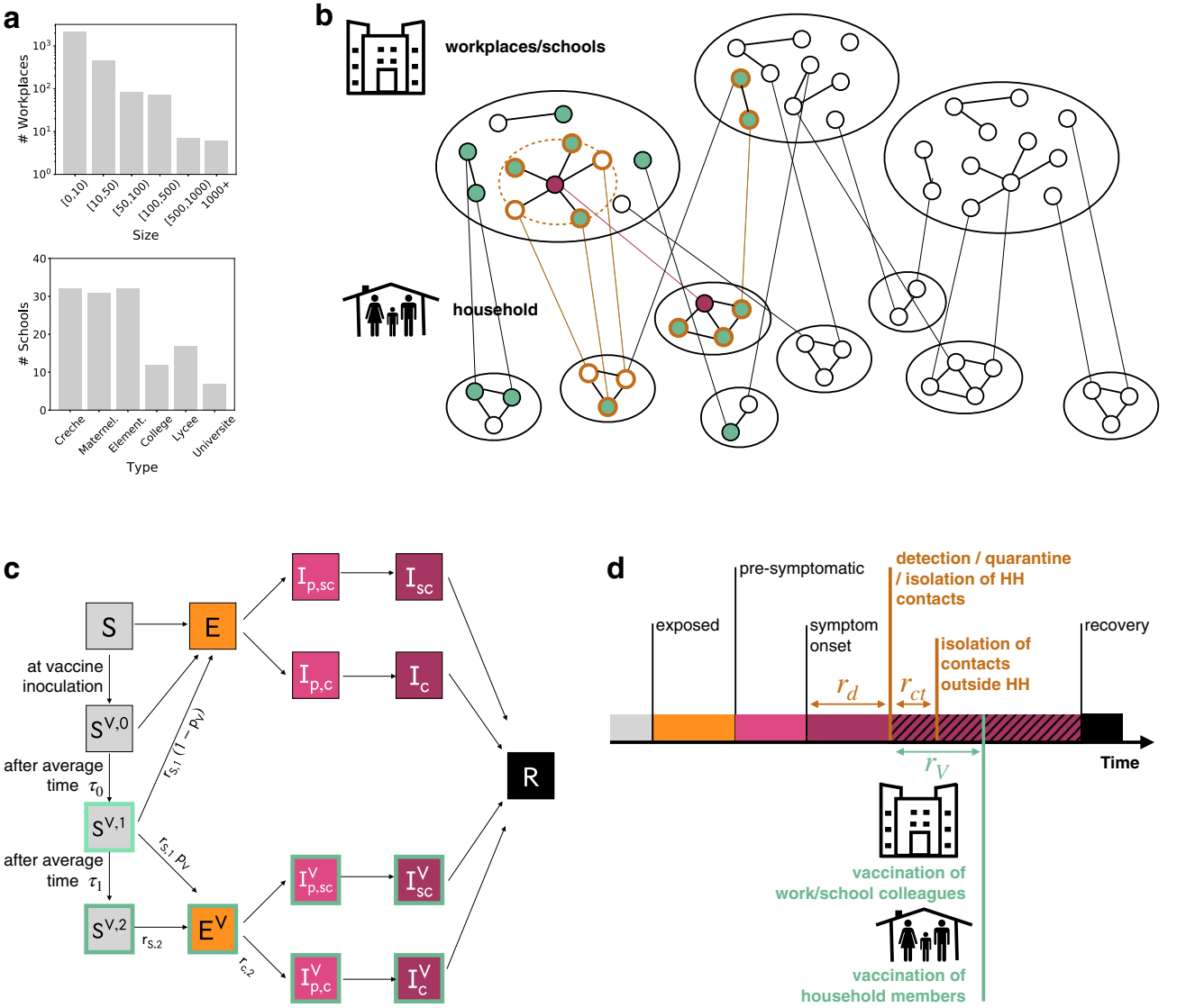


Fig. 1 Modelling reactive vaccination. a Distribution of workplace size and of school type for the municipality of Metz (Grand Est region, France), used in the simulation study. **b** Schematic representation of the population structure, the reactive vaccination and contact tracing. The synthetic population is represented as a dynamic multi-layer network, where layers encode contacts in household, workplace, school, community and transport. In the figure, school and workplace layers are collapsed and community and transport are not displayed for the sake of visualisation. Nodes repeatedly appear on both the household and the workplace/school layer. The identification of an infectious individual (in purple in the figure) triggers the detection and isolation of his/her contacts (nodes with orange border) and the vaccination of individuals attending the same workplace/school and belonging to the same household who accept to be vaccinated (green). **c** Compartmental model of COVID-19 transmission and vaccination. Description of the compartments is reported on the Methods section. **d** Timeline of events following infection for a case that is detected in a scenario with reactive vaccination. For panels **c**, **d** transition rate parameters and their values are described in the Methods and in the Supplementary Information.

In Fig. 2a–h we used the baseline parameters values presented above but varied the initial vaccination coverage level for comparison. We first considered the case of low vaccination coverage, i.e. ~30% over the population—with 15% of the [12,60] group and 90% of the 60+ group—as seen in some countries in Eastern Europe and some US counties in the fall of 2021[1,30]. For non-reactive strategies, the vaccination pace ranged between 100 and 500 first doses per 100,000 inhabitants per day. The vaccination pace in Western countries roughly fell within these extremes for the majority of the vaccination campaign, with lower values in general reached around the beginning and the end, due to delivery issues at the beginning, and difficulty in overcoming barriers to vaccination at the end[1]. For the reactive strategy, vaccine deployment is triggered by detected cases, therefore the number of doses used and the number of places where these doses are administered depends on the epidemic situation. Figure 2a shows the relative reduction in the attack rate after 2 months as a function of the number of first daily doses and Fig. 2b compares the incidence profiles under different strategies with the same number of vaccine doses. The mass, school location and workplaces/universities strategies have a similar impact on the epidemic. They lead to a reduction between 2.7% and 3% of the attack rate with 100 first doses per 100,000 inhabitants administered each day, and between 13% and 15% with 500 per 100,000 inhabitants. Among the three strategies, the reduction produced by mass vaccination was slightly lower. This is because the strategies are compared at the same number of daily vaccine doses and, in workplaces/universities and school locations, these doses were directed to a more active population—working population, or population living in large households—with a

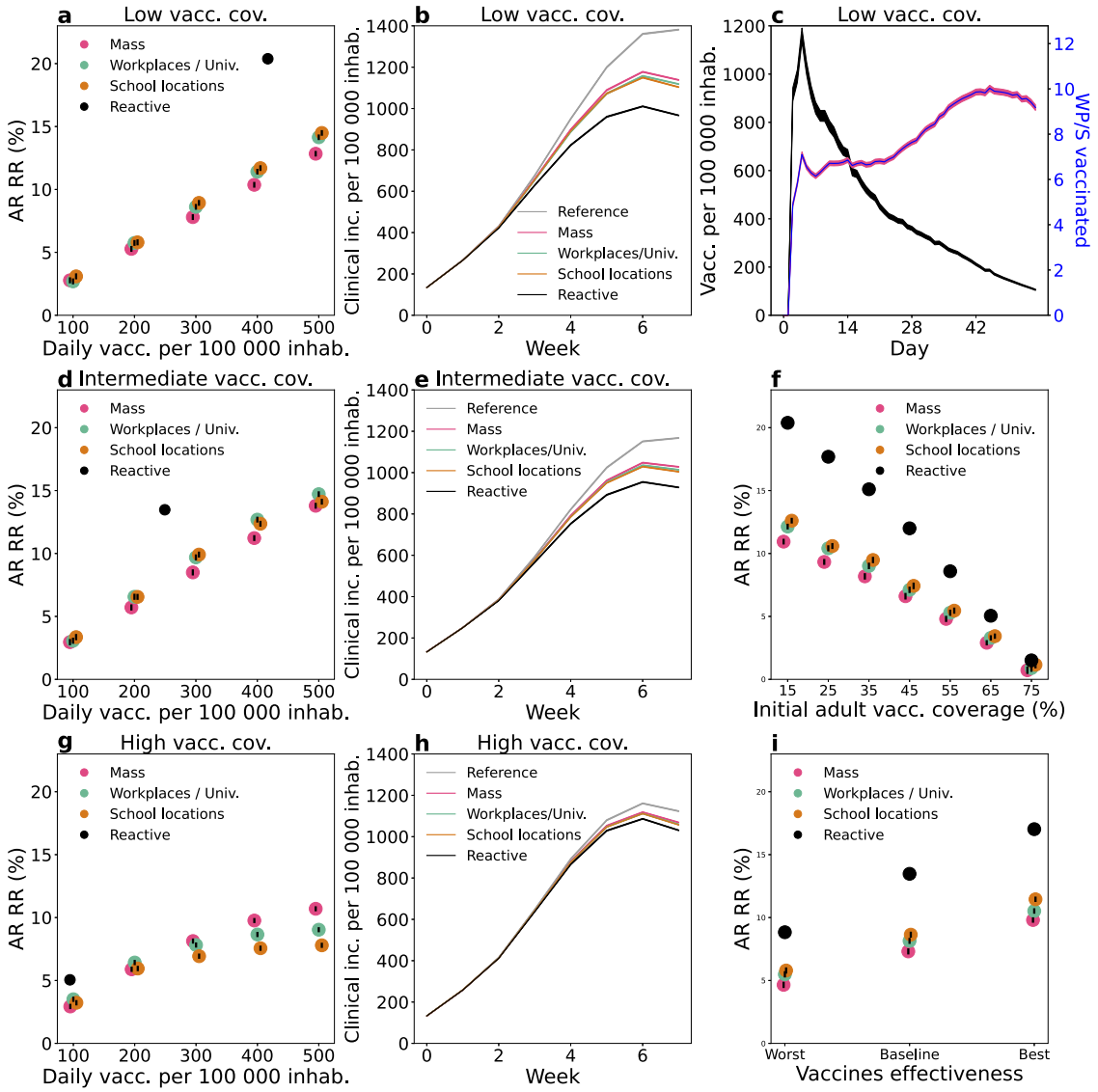


Fig. 2 Comparison between vaccination strategies. a–h Comparison between reactive and non-reactive vaccination strategies for the baseline scenario and different values of initial vaccination coverage. **a**, **d**, **g** Relative reduction (RR) in the attack rate (AR) over the first 2 months for all strategies as a function of the vaccination pace. RR is computed as $(AR_{ref} - AR)/AR_{ref}$ with AR_{ref} being the AR of the reference scenario, where no vaccination campaign is conducted during the course of the simulation and vaccination coverage remains at its initial level. AR is computed from clinical cases. Three initial vaccination coverages are investigated: 15% of adults (low) (**a**); 40% of adults (intermediate) (**d**) and 65% of adults (high) (**g**). **b**, **e**, **h** Weekly incidence of clinical cases for 100,000 inhabitants for the first 8 weeks with different vaccination strategies. The non-reactive scenarios plotted are obtained with the same average daily vaccination pace as for reactive vaccination. Low, intermediate and high vaccination coverages are investigated in **b**, **e**, **h**, respectively. **c** Number of daily first-dose vaccinations, and number of workplaces/schools (WP/S in the plot) where vaccines are deployed for the same reactive scenario as in **a**, **b**—low vacc. cov., with 15% initial vaccine coverage. **f** AR RR for different initial vaccination coverages. The four strategies are compared at equal numbers of vaccine doses. The baseline epidemic scenario of panels **a**–**h** is defined by the following key parameters: $R = 1.6$; $VE_{S,1} = 48\%$, $VE_{SP,1} = 53\%$, $VE_{S,2} = 70\%$, $VE_{SP,2} = 73\%$; initial immunity 32%; initial incidence 160 clinical cases weekly per 100,000 inhabitants; 90% of 60+ vaccinated at the beginning. **i** AR RR for different vaccine effectiveness levels, assuming intermediate vaccination coverage (40% of adults) and all other parameters as in panels **a**–**h**. The baseline vaccine effectiveness values used in the other panels is compared with a worst and a best-case scenario, defined respectively by $VE_{S,1} = 30\%$, $VE_{SP,1} = 35\%$, $VE_{S,2} = 53\%$, $VE_{SP,2} = 60\%$, and by $VE_{S,1} = 65\%$, $VE_{SP,1} = 75\%$, $VE_{S,2} = 80\%$, $VE_{SP,2} = 95\%$. For each vaccine effectiveness scenario the four strategies are compared at equal numbers of vaccine doses. In panels **a**, **d**, **f**, **g**, **i**, data are means over 2000 independent stochastic realisations and error bars are derived from the standard error of the mean—these are smaller than the size of the dot in the majority of cases. In panels **b**, **c**, **e**, **h**, continuous lines are means over 2000 independent stochastic realisations and the shaded areas are the standard error of the mean (±2SEM)—not visible in panels **b**, **e**, **h**. The distribution of outcomes over all 2000 independent stochastic realisations is provided in Supplementary Fig. 3 comparing all vaccination strategies and considering the parameterisation of Fig. 2e as an example. The following abbreviations were used in the Fig.: vacc. for vaccines, inhab. for inhabitants, cov. for coverage, inc. for incidence, univ. for universities.

greater potential to transmit the infection. The reactive vaccination produced a stronger reduction in cases than the three other strategies in the 2-months period (black dot in Fig. 2a and black line in Fig. 2b). We found that 417 doses per 100,000 inhabitants each day on average were used under the epidemic scenario considered here, yielding an attack-rate relative reduction of 20%.

In Fig. 2c we considered the same parametrisation as in Fig. 2a, b and we showed the number of first doses in time and the

number of places to vaccinate—as a proxy to the incurred logistics of a vaccine deployment. The number of daily inoculated doses was initially high, with almost 1200 doses per 100,000 inhabitants used in a day at the peak of vaccine demand, but declined rapidly afterwards down to 105 doses. The number of workplaces to vaccinate followed a different trend. It slowly increased to reach a peak and then declined. The breakdown in Supplementary Fig. 4 shows that schools and large workplaces were vaccinated at the very beginning. Thus a great number of vaccines were initially deployed in large settings, requiring many doses, while as the epidemic spread it reached a large number of small settings where only a few individuals could be vaccinated.

In Fig. 2d, e, we then considered an intermediate vaccination coverage at the beginning (40% of the [12,60] group and 90% of 60+, corresponding to ~45% of the whole population). Non-reactive strategies led to a relative reduction in the attack rate after 2 months close to that in the low initial coverage case, but the impact of reactive vaccination was reduced. Indeed, the proportion of unvaccinated people attending workplaces/schools that are targeted by vaccination is lower and fewer vaccine doses are administered, leading to a smaller impact at the population level (Fig. 2d). Still, reactive vaccination produced a 13% reduction in the attack rate using ~250 doses per 100,000 inhabitants each day on average, when the same reduction required ~400 doses per 100,000 inhabitants each day with non-reactive strategies. The impact of reactive vaccination finally became very small when initial vaccination coverage was high. Figure 2g, h shows a scenario where 65% of the [12,60] group and 90% of 60+ is vaccinated at the beginning, corresponding to ~60% of the whole population close to the coverage reached in Europe in the Fall 2021[1]. Only 94 daily vaccines per 100,000 inhabitants were used each day with a 5% reduction in the attack rate compared to a 3% reduction with non-reactive strategies for an equal number of doses. Non-reactive strategies with vaccination pace higher or equal to 300 doses per 100,000 inhabitants each day yielded a higher reduction in cases (~8% or higher).

The effect of the initial vaccination coverage on the impact of the different strategies is summarised in Fig. 2f. The relative reduction declined roughly linearly with the initial vaccination coverage. The reactive vaccination always outperformed non-reactive strategies at an equal number of doses. Nevertheless, the number of vaccinated people progressively decreased as initial vaccination coverage increased in the reactive vaccination approach, eventually reaching the point where it was less effective than non-reactive strategies with a large vaccination pace. In Fig. 2i we relaxed the baseline assumption on vaccine effectiveness and explored effectiveness parameters spanning the range of real-life estimates[17]. We found that lower vaccine effectiveness values led to a reduced effect of vaccination as expected. The difference between reactive and non-reactive strategies was also reduced.

In the Supplementary Information we compared reactive and non-reactive strategies under alternative epidemiological scenarios. In Supplementary Fig. 5 we assumed as a starting point the baseline scenario with intermediate vaccination coverage—i.e. the scenario in Fig. 2d, e with ~45% of the whole population vaccinated. We then varied key parameters, e.g. alternative values of transmission, incubation period, immunity level of the population, reduction in contacts due to social distancing, the time needed for the vaccine to become effective, compliance to vaccination and vaccine effect on the infection duration. An increase in the reproductive ratio, initial immunity and time between doses reduced the impact of the reactive vaccination. An increase in compliance to vaccination, instead, enhanced the impact of both reactive and non-reactive vaccination. Other parameters had a more limited role in strategies' effectiveness. We then considered a scenario of a flare-up of cases, as it may be caused by a new variant of concern (VOC) spreading in the territory. In Supplementary Fig. 6a–c all parameters are as the baseline case of Fig. 2d, e, except for the initial incidence. The deployment of vaccines in this case was limited and slow. We then varied other parameters, i.e. the proportion of teleworking and time from building immunity following vaccination, finding that depending on their value the reactive strategy brought limited or no benefit with respect to non-reactive strategies, when the comparison was done at an equal number of doses (Supplementary Fig. 6d, e). Eventually, we tested the robustness of our results according to the selected health outcome, using hospitalisations, ICU admissions, ICU bed occupancy, deaths, life-years lost and quality-adjusted life-years lost, finding the same qualitative behaviour (Supplementary Fig. 7).

Combined reactive and mass vaccination for managing sustained COVID-19 spread. With the high availability of vaccine doses, reactive vaccination could be deployed on top of mass vaccination. We considered the baseline scenario with intermediate vaccination coverage defined in the previous section (i.e. ~45% of the whole population vaccinated, Fig. 2d, e) and compared mass and reactive vaccination simultaneously (combined strategy) with mass vaccination alone. We focused on the first 2 months since the implementation of the vaccination strategy. At an equal number of doses within the period, the combined strategy outperformed mass vaccination in reducing the attack rate. For instance, the relative reduction in the attack rate ranged from 10%, when ~360 daily doses per 100,000 inhabitants were deployed for mass vaccination, to 16%, when the same number of doses were used for reactive and mass vaccination combined (Fig. 3a).

We explored alternative scenarios where the number of vaccines used and places vaccinated were limited due to availability and logistic constraints. We assessed the effect of three parameters: (i) the maximum daily number of vaccines that can be allocated towards reactive vaccination (with caps going from 50 to 250 per 100,000 inhabitants, compared with unlimited vaccine availability assumed in the baseline scenario), (ii) the time from the detection of a case and the vaccine deployment (set to 2 days in the baseline scenario, and here explored between 1 and 4 days) and (iii) the number of detected cases that triggers vaccination in a place (from 2 to 5 cases, vs. the baseline value of 1). The number of first-dose vaccinations in time under the different caps is plotted in Fig. 3b. A cap on the number of doses limited the impact of the reactive strategy. Figure 3c shows that the attack-rate relative reduction dropped from 16 to 6% if only a maximum of 50 first doses per 100,000 inhabitants daily was used in reactive vaccination, reaching the levels of mass vaccination only. Doubling the time required to start reactive vaccination, from 2 days to 4 days, had a limited effect on the reduction of the AR (relative reduction reduced from 16 to 15%, Fig. 3d). Increasing the number of detected cases used to trigger vaccination to 2 (respectively, 5) reduced the relative reduction to 11% (respectively, 6%) (Fig. 3e).

We so far assumed that vaccine uptake was the same in mass and reactive vaccination. This assumption is likely conservative, in that individuals may be more inclined to accept vaccination when this is proposed in the context of reactive vaccination due to the higher perceived benefit of vaccination. In Fig. 3f we departed from the baseline assumption and considered a scenario where vaccine uptake with reactive vaccination climbed to 100%. Attack-rate relative reduction increased in this case from 16 to 22%, with a demand of ~480 daily doses per 100,000 inhabitants on average.

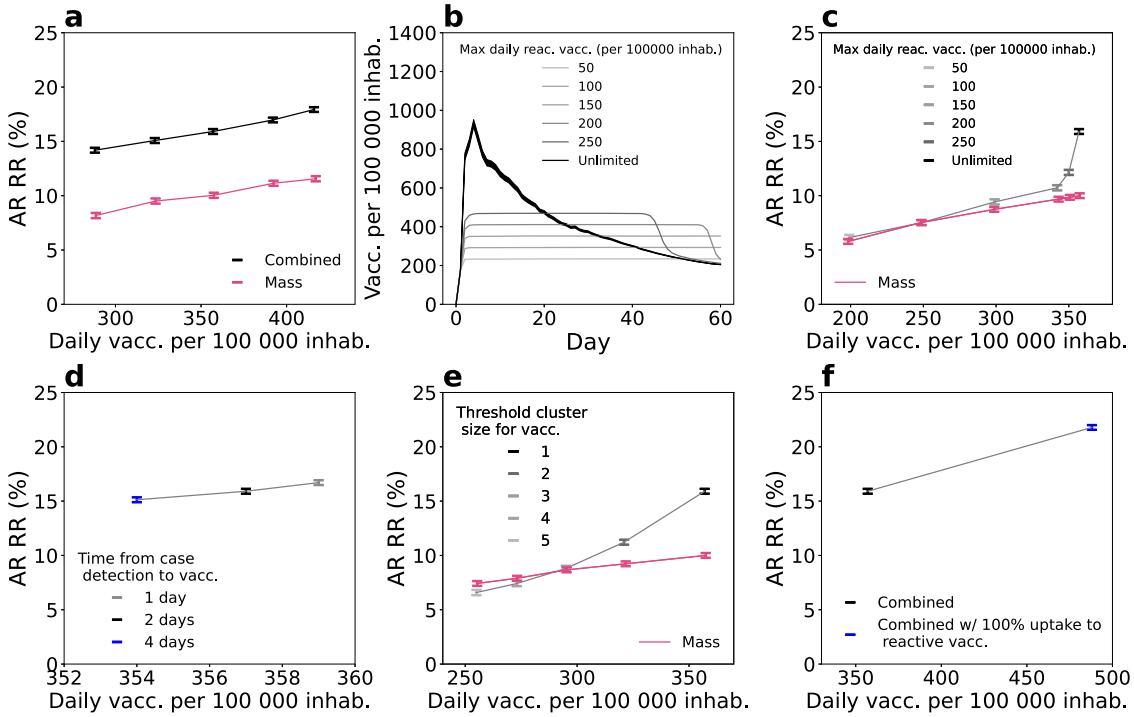


Fig. 3 Combined reactive and mass vaccination for managing sustained COVID-19 spread. a Relative reduction (RR) in the attack rate (AR) over the first 2 months for the combined strategy (mass and reactive) and the same mass strategy with the same number of first-dose vaccinations as in the combined strategy during the period. RR is computed with respect to the reference scenario with initial vaccination only, as in Fig. 2. Combined strategy is obtained by running in parallel the mass strategy—from 50 to 250 daily vaccination rate per 100,000 inhabitants—and the reactive strategy. Number of doses displayed in the *x*-axis of the figure is the total number of doses used by the combined strategy, daily. Corresponding incidence curves are reported in Supplementary Fig. 8. **b** Number of first-dose vaccinations deployed each day for the combined strategy with different daily vaccines' capacity limits. **c**, **d**, **e** AR RR for the combined strategy as a function of the average daily number of first-dose vaccinations in the 2-months period. Symbols of different colours indicate: **c** different values of daily vaccines' capacity limit; **d** different time from case detection to vaccine deployment; **e** different threshold size for the cluster to trigger vaccination. In panel **c** and **e** the curve corresponding to mass vaccination only is also plotted for comparison. **f** Comparison between 100% and baseline vaccination uptake in case of reactive vaccination. Exception for the parameters indicated in the legend we assume in all panels baseline parameter values with intermediate vaccination coverage at the beginning—i.e. $R = 1.6$; $VE_{S,1} = 48\%$, $VE_{SP,1} = 53\%$, $VE_{S,2} = 70\%$, $VE_{SP,2} = 73\%$; initial immunity 32%; initial incidence 160 clinical cases weekly per 100,000 inhabitants; vaccinated at the beginning 90% and 40% for 60+ and <60, respectively. In panels **a**, **c**, **d**–**f** data are means over 2000 independent stochastic realisations and error bars are derived from the standard error of the mean. In panel **b** continuous lines are means over 2000 independent stochastic realisations and the shaded areas are the standard error of the mean (±2SEM)—only the standard error of the unlimited case is shown for clarity. The following abbreviations were used in the Fig.: vacc. for vaccines, inhab. for inhabitants, reac. for reactive.

Combined reactive and mass vaccination for managing a COVID-19 flare-up. We previously mentioned that in a scenario of a flare-up of cases reactive vaccination would bring limited benefit compared to other strategies (Supplementary Fig. 6). Here we analyse this scenario more in-depth assuming that reactive vaccination is combined with mass vaccination but triggers an increase in vaccine uptake and is associated with enhanced TTI, as may be the case in a realistic scenario of alert due to initial VOC detection. All other parameters were as in the baseline case, with intermediate vaccination coverage at the beginning (as in Fig. 2d, e).

We assumed mass vaccination with 150 first doses per day per 100,000 inhabitants was underway from the start, as well as baseline TTI. To start a simulation, three infectious individuals carrying a VOC were introduced in the population where the virus variant was not currently circulating. Upon detection of the first case, we assumed that TTI was enhanced, finding 70% of clinical cases, 30% of subclinical cases (i.e. ~45% of all cases) and three times more contacts outside the household with 100% compliance to isolation (Supplementary Table 4)—the scenario without TTI enhancement was also explored for comparison. As soon as the number of detected cases reached a predefined threshold, reactive vaccination was started on top of the mass vaccination campaign. We assumed vaccine uptake increased to 100% for reactive vaccination but remained at its baseline value for mass vaccination.

In Fig. 4 we compare the combined scenario with mass vaccination alone at an equal number of doses and investigate starting reactive vaccination after 1, 5 or 10 detected cases. With reactive vaccination starting from the first detected case, the attack rate decreased by ~10%, compared with the mass scenario. However, the added value of reactive vaccination decreased if the start of the intervention was delayed. Without enhancement in TTI and increase in vaccine uptake, attack-rate values were much higher and the benefit of reactive vaccination over the mass vaccination was lower (~3%).

In Supplementary Fig. 10 we show different epidemic scenarios, testing different values for the transmissibility and vaccine effectiveness—including worst-case vaccine effectiveness, and R as high as 1.8—and found similar trends. Finally, we analysed the impact of vaccination on the flare-up extinction (Supplementary Fig. 11). With the parameterisation of Fig. 4a, the probabilities of extinction were 5.5% and 6% with mass and combined strategies, respectively. These values increased to

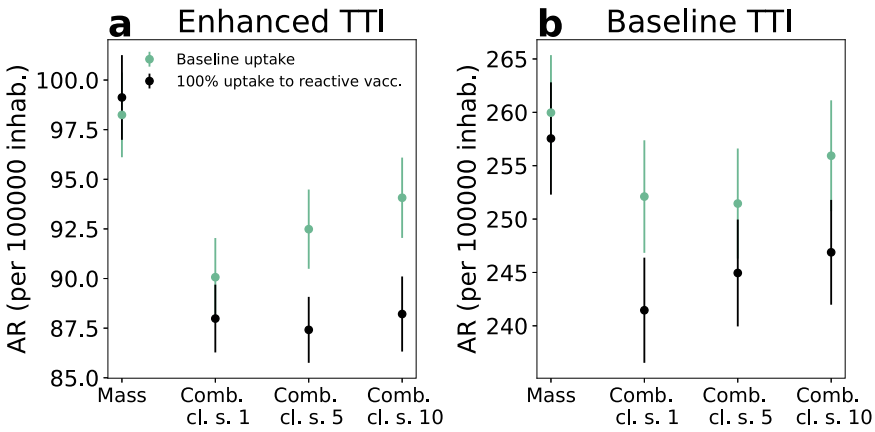


Fig. 4 Combined reactive and mass vaccination for managing a COVID-19 flare-up. a, b Attack rate (AR) per 100,000 inhabitants in the first 2 months for the enhanced (**a**) and baseline (**b**) TTI scenarios described in the main text. Four vaccination strategies are compared: mass only, combined where the reactive vaccination starts at the detection of 1, 5, 10 cases (Comb. cl. s. = 1, 5, 10 in the Fig.). For mass vaccination the number of first-dose vaccinations during the period is the same as in the comb. cl. s = 1 of the same scenario. Except when otherwise indicated parameters are the ones of the baseline epidemic scenario with intermediate vaccination coverage at the beginning—i.e. $R = 1.6$; $VE_{S,1} = 48\%$, $VE_{SP,1} = 53\%$, $VE_{S,2} = 70\%$, $VE_{SP,2} = 73\%$; initial immunity 32%; vaccinated at the beginning 90% and 40% for 60+ and <60, respectively. In both panels, data are means over 8000 independent stochastic realisations and error bars are the standard error of the mean (±2SEM). Corresponding incidence curves are reported in Supplementary Fig. 9. The following abbreviations were used in the Fig.: vacc. for vaccines, inhab. for inhabitants.

15% and 18% in a best-case scenario where vaccine protection occurred sooner after the first dose, vaccine efficacy was larger and TTI further strengthened.

Discussion

The rapid rise of more transmissible SARS-CoV-2 variants has made the course of the COVID-19 pandemic unpredictable, posing a persistent public health threat that jeopardises the return to normal life[25,31–35]. More transmissible viruses call for vaccination of a larger portion of the population and increased accessibility and adaptation to a rapidly changing epidemic situation[2]. In this context, we analysed reactive vaccination of workplaces, universities and schools to assess its potential role in managing the epidemic.

The agent-based model used in this study accounted for the major factors affecting the effectiveness of reactive vaccination: disease natural history, vaccine characteristics, individual contact behaviour and logistical constraints. For a broad range of epidemic scenarios, reactive vaccination reduced the spread of COVID-19 more than non-reactive vaccination strategies—including untargeted mass vaccination—for an equal number of doses used over a period of 2 months. In addition, combining reactive and mass vaccination was more effective than mass vaccination alone. For instance, in a scenario of moderate/high incidence starting with 45% vaccination coverage, we found that the relative reduction in the attack rate over 2 months increased from 10 to 16% if 350 first vaccine doses per 100,000 habitants per day were used in a combined mass/reactive vaccination approach instead of mass only. However, the advantage of reactive vaccination was limited or nought with respect to non-reactive strategies under certain circumstances, as the number of doses administered with the reactive vaccination depended on the number and pace of occurrence and detection of COVID-19 cases. This occurred when vaccination coverage was already high at the beginning and only a few people could be vaccinated around detected cases, or in a flare-up scenario when only a few cases were detected. Non-reactive strategies could then be more effective provided the pace of vaccine administration was large enough. But even in these situations, adding reactive vaccination to mass vaccination could become of interest again by triggering an increase in vaccine uptake, all the more when combined with enhanced TTI.

Reactive vaccination has been studied for smallpox, cholera and measles, among others[5–7,36,37]. Hotspot vaccination was found to help in cholera outbreak response in both modelling studies and outbreak investigation[37,38]. It may target geographic areas defined at spatial resolution as diverse as districts within a country, or neighbourhoods within a city, according to the situation. For Ebola and smallpox, ring vaccination was successfully adopted to accelerate epidemic containment[5–7]. These infections, though, have features making the approach a priori sensible: vaccine-induced immunity mounts rapidly compared to the incubation period and the mere absence of pre-symptomatic and asymptomatic transmission makes it possible to reach secondary cases before they start transmitting. Ring vaccination is also relevant when the vaccine has post-exposure effects[8]. Reactive vaccination in schools and university campuses has been implemented in the past to contain outbreaks of meningitis[39] and measles[40,41].

For COVID-19, the use of reactive vaccination has been reported in Ontario, the UK, Germany and France among others[2,42–47]. In these places, vaccines were directed to communities, neighbourhoods or building complexes with a large number of infections or presenting epidemic clusters or surge of cases due to virus variants. While the goal of these campaigns was to minimise the spread of the virus, it also addressed inequalities in access to healthcare and increased fairness, since a surge of cases may happen where people have difficulty in isolating due to poverty and house crowding[48]. In France, reactive vaccination was implemented to contain the emergence of variants of concern in the municipalities of Bordeaux, Strasbourg and Brest[45–47]. In the municipality of Strasbourg, vaccination slots dedicated to students were created following the identification of a Delta cluster in an art school[45]. Despite the interest in the strategy and its inclusion in the COVID-19 response plans, very limited work has been done so far to quantify its effectiveness[49,50]. A modelling study on ring vaccination suggested that the strategy could be valuable if the vaccine had post-exposure efficacy and a large proportion of contacts could be identified[50]. Still, post-exposure effects of the vaccine remain currently hypothetical[51], and it is likely that the vaccination of

the first ring of contacts alone would bring little benefit, if at all. We have here tested reactive vaccination of workplaces and schools, since focusing on these settings may be an efficient way to easily reach an extended group of contacts. Workplaces have been found to be an important setting for COVID-19 transmission, especially specific workplaces where conditions are more favourable for spreading[52,53]. The university setting also plays a central role in COVID-19 transmission, due to the high number of contacts among students, particularly if sharing common spaces in residence accommodations[54]. Model results show that reactive vaccination of these settings could have in many circumstances a stronger impact than simply reinforcing vaccination in these settings.

Importantly, the effectiveness of the reactive strategy depends on the epidemic context. We found that the higher the overall vaccination coverage, the less reactive vaccination would be of interest compared to non-reactive strategies if it did not increase vaccine uptake. For example, with >40% vaccination coverage among adults, the relative reduction in attack rates with reactive vaccination is smaller than with non-reactive alternatives provided large enough vaccination pace and no increase in vaccine uptake with reactive vaccination. Indeed, with large vaccination coverage, only a few individuals who have not been vaccinated before can be reactively proposed for vaccination, leading to few shots in case of vaccine hesitancy. Moreover, the detection of clusters is more difficult in a highly vaccinated population, where breakthrough infections in the vaccinated yield a large proportion of subclinical cases that are harder to detect.

If initial vaccination coverage is not too high, the feasibility and advantage of the inclusion of reactive vaccination imply a trade-off between epidemic intensity and logistical constraints. At a moderate/high incidence level, combining reactive and mass vaccination would substantially decrease the attack rate compared to mass vaccination for the same number of doses, but the large initial demand in vaccines may exceed the available stockpiles or capacity to deliver. The timely deployment of additional personnel in mobile vaccine units and the need to quickly inform the population by communication campaigns is indeed key to guarantee the success of the campaign. We explored with the model the key variables that would impact the strategy effectiveness. Delaying the deployment of vaccines in workplaces/schools upon the detection of a case (from 2 to 4 days on average) would not have a strong impact on its effectiveness. However, vaccines should be deployed at the detection of the first case to avoid substantially limiting the impact of the strategy—e.g. the relative reduction went from 16 to 6% when workplaces/schools were only vaccinated after the detection of 5 cases (Fig. 3e).

In the case of a COVID-19 flare-up the reactive strategy may bring an advantage if the reactive strategy starts early, is combined with increased TTI and triggers an increase in vaccine uptake. Starting early after the introduction of a first VOC case requires that tests for the detection of variants must be carried out regularly and with large coverage. Genomic surveillance has ramped up in many countries since the emergence of the Alpha variant in late 2020. For example, as of Fall 2021, nationwide surveys are conducted almost weekly in France to fully sequence the viral genome in randomly selected positive samples[55]. A proportion of positive tests are also routinely screened for key mutations to monitor the circulation of the main variants registered as VOC or VUI[55]. This surveillance protocol contributes to quickly identifying the presence of variants, but does not guarantee that interventions start with the few first cases, even more as the relaxation of social restrictions may lead to super spreading events. A strong intensification of TTI[23] must also be part of the wider response plan including reactive vaccination. Rapid and efficient TTI efficiently mitigates spread on its own, but it is also instrumental to the success of reactive vaccination by triggering vaccination in households, workplaces and schools. Last, an increased level of vaccine uptake is essential for reactive vaccination to be of interest. Vaccination coverage remains highly heterogeneous worldwide and, as of Fall 2021, low in many countries of Eastern Europe and in many counties in the US[1,30]. Besides the individuals who oppose vaccination, a reactive strategy combined with the presence of a VOC may help increasing the acceptability of the vaccine by making it more accessible and anticipating an immediate benefit against the risk of infection. An increase in vaccine uptake was indeed observed in the context of a reactive vaccination campaign during the course of a measles outbreak[4]. Reactive vaccination could therefore be a means to improve access and acceptability in case of a flare-up.

The study is affected by several limitations. First, the synthetic population used in the study accounts for the repartition of contacts across workplaces, schools, households, etc., informed by contact surveys. However, number of contacts and risk of transmission may vary greatly according to the kind of occupation. The synthetic population accounts for this variability assuming that the average number of contacts from one workplace to another is gamma distributed[10], but no data were available to inform the model in this respect. Second, we model vaccination uptake according to age only, when it is determined by several socio-demographic factors. Clusters of vaccine-hesitant individuals may play an important role in the dynamics and facilitate the epidemic persistence in the population, as it is described for measles[56]. In those countries where vaccination coverage is high, heterogeneities in attitude toward vaccination may have an impact. Third, the agent-based model is calibrated from French socio-demographic data. The results of this study can be extended to countries with similar societal structure and contact patterns, e.g. other developed countries[57]. Still, COVID-19 transmission potential, level of disease-induced immunity, vaccination coverage, and extent of social restrictions vary substantially from one country to another. In addition, the waning of immunity since vaccination and recommendations for booster doses affect the level of protection of the population already vaccinated and consequently the impact of reactive vaccination. The large set of scenarios explored and reported in the Supplementary Information is intended to fully understand the interplay between epidemic spread and reactive vaccination and aid planning in case of future epidemic surges.

Methods

Synthetic population. We used a synthetic population for a French municipality based on the National Institute of Statistics and Economic Studies (INSEE) censuses and French contact survey information[10,58]. This included the following input files: (i) a setting-specific, time-varying network of daily face-to-face contacts; (ii) the correspondence between individuals and their age, (iii) between individuals and the household they belong to, (iv) between individuals and their school, (v) and between individuals and their workplace. The synthetic population has an age pyramid, household composition, number of workplaces by size and number of schools by type, reproducing INSEE statistics. Daily face-to-face contacts among individuals are labelled according to the setting in which they occur (either household, workplace, school, community or transport) and they have assigned a daily frequency of activation, to explicitly model recurrent and sporadic contacts. We considered the municipality of Metz in the Grand Est region, which has 117,492 inhabitants, 131 schools (from kindergarten to University) and 2888 workplaces (Fig. 1a). A detailed description of how the population was generated is provided in[10]. Information about how to access population files is provided in the Data availability section.

Overview of the model. The model was written in C/C++, and is stochastic and discrete-time. It accounts for the following components: (i) teleworking and social distancing, (ii) COVID-19 transmission, accounting for the effect of the vaccine; (iii) test-trace-isolate; (iv) vaccine deployment. Model output included time series of incidence (clinical and subclinical cases), detailed information on infected cases (time of infection, age, vaccination status), vaccines administered according to the strategy, number of workplaces where vaccines are deployed. Different epidemic

scenarios were explored and compared. In the Supplementary Information we also analysed hospitalisation entries, deaths, ICU entries, life-year lost, quality-adjusted life-year, ICU bed occupancy. These quantities were computed by postprocessing output files containing detailed information on infected cases.

Teleworking and social distancing. Teleworking and other social restrictions may alter the repartition of contacts across settings and in turn the effectiveness of vaccination strategies[23]. We thus explicitly accounted for this ingredient in the model. Specifically, to model teleworking we assumed a proportion of individuals were absent from work, modelled by erasing working contacts and transport contacts of these individuals. To account for the reduction in social encounters due to the closure of restaurants and other leisure activities we removed a proportion of contacts from the community layer. In Western countries, the level of restrictions varied greatly both by country and in the time since vaccines were first deployed at the beginning of 2021. We set the contact reduction in the community to 5% and the teleworking to 10%. These were close to the reduction values reported by google mobility reports for France during Autumn 2021[59], and fell within the range of European countries' estimates. Note that levels of teleworking ~10% for European countries were reported also by other sources[60]. Scenarios with different levels of contact reduction were compared in the Supplementary Information. Telework and social distancing were implemented at the beginning of the simulation and remained constant for the duration of the simulation. Importantly, the reproductive ratio was set to the desired value, independently by the level of contact reduction, as described in the Supplementary Information.

COVID-19 transmission model. We used an extension of the transmission model in ref. [10] (see Fig. 1c). This accounted for heterogeneous susceptibility and severity across age groups[61,62], the presence of an exposed and a pre-symptomatic stage[9] and two different levels of infection outcome—subclinical, corresponding to asymptomatic or paucisymptomatic infection and clinical, corresponding to moderate to critical infection[61,63]. Precisely, susceptible individuals, if in contact with infectious ones, could get infected and enter the exposed compartment (E). After an average latency period ϵ^{-1} they became infectious, developing a subclinical infection (I_{sc}) with age-dependent probability p_{sc}^{A} and a clinical infection (I_c) otherwise. From E, before entering either I_{sc} or I_c, individuals entered first a prodromal phase (either $I_{p,sc}$ or $I_{p,c}$, respectively), that lasted on average μ_p^{-1} days. Compared to $I_{p,c}$ and I_c individuals, individuals in the $I_{p,sc}$ and I_{sc} compartments had reduced transmissibility rescaled by a factor β_I. With rate μ infected individuals became recovered. Age-dependent susceptibility and age-dependant probability of clinical symptoms were parametrised from[61]. In addition, transmission depended on setting as in[10]. We assumed that the time spent in the E, $I_{p,sc}$ and $I_{p,c}$ was Erlang distributed with shape 2, and rate 2ϵ for E, and $2\mu_p$ for $I_{p,sc}$ and $I_{p,c}$. Time spent in I_{sc} and I_c was exponentially distributed. Parameters and their values are summarised in Supplementary Table 1.

We modelled vaccination with a leaky vaccine, partially reducing both the risk of infection (i.e. reduction in susceptibility, VE_S) and infection-confirmed symptomatic illness (VE_{SP})[15]. The level of protection increased progressively after the inoculation of the first dose. In our model we did not explicitly account for the two-dose administration, but we accounted for two levels of protection—e.g. a first one approximately in between the two doses and a second one after the second dose. Vaccine efficacy was zero immediately after inoculation, mounting then to an intermediate level ($VE_{S,1}$ and $VE_{SP,1}$) and a maximum level later ($VE_{S,2}$ and $VE_{SP,2}$). This is represented through the compartmental model in Fig. 1c. Upon administering the first dose, S individuals became, $S^{V,0}$, i.e. individuals that are vaccinated, but have no vaccine protection. If they did not become infected, they entered stage $S^{V,1}$, where they were partially protected, then stage $S^{V,2}$ where vaccine protection was maximum. Time spent in $S^{V,0}$ and $S^{V,1}$ was Erlang distributed with shape 2 and rate $2/\tau_0$ and $2/\tau_1$ for $S^{V,0}$ and $S^{V,1}$, respectively. $S^{V,1}$ and $S^{V,2}$ individuals had reduced probability of getting infected by a factor $r_{S,1} = (1 - VE_{S,1})$ and $r_{S,2} = (1 - VE_{S,2})$, respectively. In case of infection, $S^{V,2}$ individuals progressed first to exposed vaccinated (E^V), then to either preclinical or pre-subclinical vaccinated ($I_{p,c}^V$ or $I_{p,sc}^V$) that were followed by clinical and subclinical vaccinated, respectively (I_c^V or I_{sc}^V). Probability of becoming $I_{p,c}^V$ from E^V was reduced of a factor $r_{c,2} = (1 - VE_{SP,2})(1 - VE_{S,2})^{-1}$. For the $S^{V,1}$ individuals that get infected we assumed a polarised vaccine effect, i.e. they can enter either in E^V, with probability p_V, or in E (Fig. 1c). The value of p_V was set based on $VE_{SP,1}$ through the relation $(1 - VE_{SP,1}) = (1 - VE_{S,1})(p_V r_{c,2} + (1 - p_V))$. We assumed no reduction in infectiousness for vaccinated individuals. However, we accounted for a 25% reduction in the duration of the infectious period as reported in refs. [64,65].

Under the assumption that no serological/virological/antigenic test is done before vaccine administration, the vaccine was administered to all individuals, except for clinical cases who showed clear signs of the disease or individuals that were detected as infected by the TTI in place. In our model a vaccine administered to infected or recovered individuals had no effect.

In the baseline scenario we parametrised $VE_{SP,1}$, $VE_{SP,2}$, $VE_{S,1}$ and $VE_{S,2}$ by taking values in the middle of estimates reported in the systematic review by Higdon and collaborators[17] for the Comirnaty vaccine and the Delta variant[66]. Chosen values of $VE_{SP,2}$, and $VE_{S,2}$ are also comparable with the effectiveness estimates reported in a meta-analysis for the Delta variant, complete vaccination, all vaccines combined[18]. We also tested values on the upper and lower extremes of the range of estimates of[17]. Parameters are listed in Supplementary Table 2.

Test-trace-isolate. We model a baseline TTI accounting for case detection, household isolation and manual contact tracing. Fifty percent of individuals with clinical symptoms were assumed to get tested and to isolate if positive. We assumed an exponentially distributed delay from symptoms onset to case detection and its isolation with 3.6 days on average. Once a case was detected, his/her household members isolated with probability $p_{ct,HH}$, while other contacts isolated with probabilities $p_{ct,A}$ and $p_{ct,Oth}$, for acquaintances and sporadic contacts, respectively. In addition to the detection of clinical cases, we assumed that a proportion of subclinical cases were also identified (10%). Isolated individuals resumed normal daily life after 10 days unless they still had clinical symptoms after the time had passed. They could, however, decide to drop out from isolation each day with a probability of 13% if they did not have symptoms[67].

In the scenario of virus re-introduction we considered enhanced TTI, corresponding to a situation of case investigation, screening campaign and sensibilisation (prompting higher compliance to isolation). We assumed a higher detection of clinical and subclinical cases (70% and 30%, respectively), perfect compliance to isolation by the index case and household members and a three-fold increase in contacts identified outside the household.

A step-by-step description of contact tracing is provided in the Supplementary Information. Parameters for baseline TTI are provided in Supplementary Table 3, while parameters for enhanced TTI are provided in Supplementary Table 4.

Vaccination strategies. A vaccine opinion (willingness or not to vaccinate) was stochastically assigned to each individual at the beginning of the simulation depending on age (below/above 65 years old). The opinion did not change during the simulation. In some scenarios we assumed that all individuals were willing to accept the vaccine in case of reactive vaccination, while maintaining the opinion originally assigned to them when the vaccine was proposed in the context of non-reactive vaccination. Only individuals above a threshold age, $a_{th,V} = 12$ years old, were vaccinated. We assumed that a certain fraction of individuals were vaccinated at the beginning of the simulation according to the age group ([12,60], 60+). We compared the following vaccination strategies:

Mass. V_{daily} randomly selected individuals were vaccinated each day until a V_{tot} limit was reached.

Workplaces/universities. Random workplaces/universities were selected each day. All individuals belonging to the place, willing to be vaccinated, and not isolated at home that day were vaccinated. Individuals in workplaces/universities were vaccinated each day until the daily limit, V_{daily}, was reached. No more than V_{tot} individuals were vaccinated during the course of the simulation. We assumed that only workplaces with $\text{size}_{\text{th}} = 20$ employees or larger implemented vaccination.

School location. Random schools, other than universities, were selected each day and a vaccination campaign was conducted in the places open to all household members of school students. All household members willing to be vaccinated, above the threshold age and not isolated at home that day were vaccinated. No more than V_{daily} individuals were vaccinated each day and no more than V_{tot} individuals were vaccinated during the course of the simulation.

Reactive. When a case was detected, vaccination was done in her/his household with rate r_V. When a cluster—i.e. at least n_{cl} cases detected within a time window of length T_{cl}—was detected in a workplace/school, vaccination was done in that place with rate r_V. In the baseline scenario, we assumed vaccination in the workplace/school was triggered by one single infected individual ($n_{cl} = 1$). In both households and workplaces/schools, all individuals belonging to the place above the threshold age and willing to be vaccinated were vaccinated. Individuals that were already detected and isolated at home were not vaccinated. No more than V_{daily} individuals were vaccinated each day and no more than V_{tot} individuals were vaccinated during the course of the simulation. In the baseline scenario these quantities were unlimited, i.e. all individuals to be vaccinated in the context of reactive vaccination were vaccinated.

Parameters and their values are summarised in Supplementary Table 5.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The synthetic population used in the analysis is available on zenodo[68].

Code availability

We provide all C/C++ code files of the model on zenodo[68].

Received: 26 July 2021; Accepted: 17 February 2022;
Published online: 17 March 2022

References

1. Ritchie, H. et al. Our World In Data. Coronavirus Pandemic (COVID-19). https://ourworldindata.org/coronavirus (2021).
2. European Centre for Disease Prevention and Control. *Overview of the implementation of COVID-19 vaccination strategies and deployment plans in the EU/EEA*. https://www.ecdc.europa.eu/en/publications-data/overview-implementation-covid-19-vaccination-strategies-and-deployment-plans (2021).
3. STAT. *To vaccinate more Americans, lean into outbreaks*. https://www.statnews.com/2021/08/19/lean-into-outbreaks-to-vaccinate-more-americans/ (2021).
4. Le Menach, A. et al. Increased measles–mumps–rubella (MMR) vaccine uptake in the context of a targeted immunisation campaign during a measles outbreak in a vaccine-reluctant community in England. *Vaccine* **32**, 1147–1152 (2014).
5. Merler, S. et al. Containing Ebola at the source with Ring vaccination. *PLOS Neglected Tropical Dis.* **10**, e0005093 (2016).
6. Henao-Restrepo, A. M. et al. Efficacy and effectiveness of an rVSV-vectored vaccine expressing Ebola surface glycoprotein: interim results from the Guinea ring vaccination cluster-randomised trial. *Lancet* https://doi.org/10.1016/S0140-6736(15)61117-5 (2015).
7. Geddes, A. M. The history of smallpox. *Clin. Dermatol.* **24**, 152–157 (2006).
8. Gallagher, T. & Lipsitch, M. Postexposure effects of vaccines on infectious diseases. *Epidemiol. Rev.* **41**, 13–27 (2019).
9. He, X. et al. Temporal dynamics in viral shedding and transmissibility of COVID-19. *Nat. Med.* https://doi.org/10.1038/s41591-020-0869-5 (2020).
10. Moreno López, J. A. et al. Anatomy of digital contact tracing: role of age, transmission setting, adoption and case detection. *Sci. Adv.* https://doi.org/10.1126/sciadv.abd8750 (2021).
11. Ajelli, M. et al. Comparing large-scale computational approaches to epidemic modeling: agent-based versus structured metapopulation models. *BMC Infect. Dis.* **10**, 190 (2010).
12. Chao, D. L., Halloran, M. E., Obenchain, V. J. & Longini, I. M. Jr FluTE, a publicly available stochastic influenza epidemic simulation model. *PLoS Comput. Biol.* **6**, e1000656 (2010).
13. Willem, L., Verelst, F., Bilcke, J., Hens, N. & Beutels, P. Lessons from a decade of individual-based models for infectious disease transmission: a systematic review (2006-2015). *BMC Infect. Dis.* **17**, 612 (2017).
14. Liu, Q.-H. et al. Measurability of the epidemic reproduction number in data-driven contact networks. *PNAS* **115**, 12680–12685 (2018).
15. Basta, N. E., Halloran, M. E., Matrajt, L. & Longini, I. M. Estimating influenza vaccine efficacy from challenge and community-based study data. *Am. J. Epidemiol.* **168**, 1343–1352 (2008).
16. Polack, F. P. et al. Safety and efficacy of the BNT162b2 mRNA Covid-19 vaccine. *N. Engl. J. Med.* **383**, 2603–2615 (2020).
17. Higdon, M. M. et al. A systematic review of COVID-19 vaccine efficacy and effectiveness against SARS-CoV-2 infection and disease. *medRxiv* https://www.medrxiv.org/content/10.1101/2021.09.17.21263549v1 (2021).
18. Harder, T. et al. Effectiveness of COVID-19 vaccines against SARS-CoV-2 infection with the Delta (B.1.617.2) variant: second interim results of a living systematic review and meta-analysis, 1 January to 25 August 2021. *Euro. Surveill.* **26**, 2100920 (2021).
19. Institut Pasteur. *Proportion de la population infectée par SARS-CoV-2*. https://modelisation-covid19.pasteur.fr/realtime-analysis/infected-population/ (2021).
20. Hozé, N. et al. Monitoring the proportion of the population infected by SARS-CoV-2 using age-stratified hospitalisation and serological data: a modelling study. *Lancet Public Health* **6**, e408–e415 (2021).
21. Krymova, E. et al. Trend estimation and short-term forecasting of COVID-19 cases and deaths worldwide. *arXiv* https://arxiv.org/abs/2106.10203 (2021).
22. SPF. *COVID-19: point épidémiologique du 6 mai 2021*. https://www.santepubliquefrance.fr/maladies-et-traumatismes/maladies-et-infections-respiratoires/infection-a-coronavirus/documents/bulletin-national/covid-19-point-epidemiologique-du-6-mai-2021 (2021).
23. Pullano, G. et al. Underdetection of COVID-19 cases in France threatens epidemic control. *Nature* https://doi.org/10.1038/s41586-020-03095-6 (2020).
24. YouGov. *COVID-19: Willingness to be vaccinated*. https://yougov.co.uk/topics/international/articles-reports/2021/01/12/covid-19-willingness-be-vaccinated (2021).
25. Tran Kiem, C. et al. A modelling study investigating short and medium-term challenges for COVID-19 vaccination: From prioritisation to the relaxation of measures. *EClinicalMedicine*. **38**, 101001 (2021).
26. Bubar, K. M. et al. Model-informed COVID-19 vaccine prioritization strategies by age and serostatus. *Science* https://doi.org/10.1126/science.abe6959 (2021).
27. Matrajt, L., Eaton, J., Leung, T. & Brown, E. R. Vaccine optimization for COVID-19: Who to vaccinate first? *Sci. Adv.* **7**, eabf1374 (2020).
28. Lu, D. et al. Data-driven estimate of SARS-CoV-2 herd immunity threshold in populations with individual contact pattern variations. *medRxiv* https://doi.org/10.1101/2021.03.19.21253974 (2021).
29. Marziano, V. et al. Parental vaccination to reduce measles immunity gaps in Italy. *eLife* **8**, e44942 (2019).
30. Bansal Lab. *US COVID-19 Vaccination Tracking*. http://www.vaccinetracking.us (2021).
31. Huang, B. et al. Integrated vaccination and physical distancing interventions to prevent future COVID-19 waves in Chinese cities. *Nat. Hum. Behav.* **5**, 695–705 (2021).
32. Moore, S., Hill, E. M., Tildesley, M. J., Dyson, L. & Keeling, M. J. Vaccination and non-pharmaceutical interventions for COVID-19: a mathematical modelling study. *Lancet Infect. Dis.* **21**, 793–802 (2021).
33. Yang, J. et al. Despite vaccination, China needs non-pharmaceutical interventions to prevent widespread outbreaks of COVID-19 in 2021. *Nat. Hum. Behav.* https://doi.org/10.1038/s41562-021-01155-z (2021).
34. Gozzi, N., Bajardi, P. & Perra, N. The importance of non-pharmaceutical interventions during the COVID-19 vaccine rollout. *PLoS Comput. Biol.* **17**, e1009346 (2021).
35. Bosetti, P. et al. Epidemiology and control of SARS-CoV-2 epidemics in partially vaccinated populations: a modeling study applied to France. *BMC Med* **20**, 33 (2022).
36. Grais, R. F. et al. Time is of the essence: exploring a measles outbreak response vaccination in Niamey, Niger. *J. R. Soc. Interface* https://doi.org/10.1098/rsif.2007.1038 (2007).
37. Finger, F. et al. The potential impact of case-area targeted interventions in response to cholera outbreaks: a modeling study. *PLoS Med.* **15**, e1002509 (2018).
38. Martin, S. et al. Post-licensure deployment of oral cholera vaccines: a systematic review. *Bull. World Health Organ* **92**, 881–893 (2014).
39. Capitano, B., Dillon, K., LeDuc, A., Atkinson, B. & Burman, C. Experience implementing a university-based mass immunization program in response to a meningococcal B outbreak. *Hum. Vaccines Immunotherapeutics* **15**, 717–724 (2019).
40. Pegorie, M. et al. Measles outbreak in Greater Manchester, England, October 2012 to September 2013: epidemiology and control. *Euro. Surveill.* **19**, 20982 (2014).
41. Kirolos, A. et al. Imported case of measles in a university setting leading to an outbreak of measles in Edinburgh, Scotland from September to December 2016. *Epidemiol. Infect.* **146**, 741–746 (2018).
42. COVID-19 (coronavirus) in Ontario. *Ontario's COVID-19 vaccination plan*. https://covid-19.ontario.ca/ontarios-covid-19-vaccination-plan (2021).
43. BBC News. *Covid-19: More variant hotspots to get surge tests and jabs*. https://www.bbc.com/news/uk-57172139 (2021).
44. DW.COM. *COVID: Cologne project aims to vaccinate urban hot spots*. https://www.dw.com/en/covid-cologne-project-aims-to-vaccinate-urban-hot-spots/a-57472989 (2021).
45. ARS Grand Est. *COVID-19: présence du variant indien dans l'Eurométropole et plan d'actions immédiat*. http://www.grand-est.ars.sante.fr/covid-19-presence-du-variant-indien-dans-leurometropole-et-plan-dactions-immediat (2021).
46. ARS Bretagne. *COVID-19: en Pays de Brest, la vaccination s'accélère*. http://www.bretagne.ars.sante.fr/covid-19-en-pays-de-brest-la-vaccination-saccelere (2021).
47. ARS Nouvelle Aquitaine. *Communiqué de presse - Covid-19 - La nécessité de se faire vacciner rapidement pour éviter la propagation du virus, notamment du variant delta*. http://www.nouvelle-aquitaine.ars.sante.fr/communique-de-presse-covid-19-la-necessite-de-se-faire-vacciner-rapidement-pour-eviter-la (2021).
48. Valdano, E., Lee, J., Bansal, S., Rubrichi, S. & Colizza, V. Highlighting socio-economic constraints on mobility reductions during COVID-19 restrictions in France can inform effective and equitable pandemic response. *J. Travel Med.* **28**, taab045 (2021).
49. Xu, W., Su, S. & Jiang, S. Ring vaccination of COVID-19 vaccines in medium- and high-risk areas of countries with low incidence of SARS-CoV-2 infection. *Clin. Transl. Med.* **11**, e331 (2021).
50. MacIntyre, C. R., Costantino, V. & Trent, M. Modelling of COVID-19 vaccination strategies and herd immunity, in scenarios of limited and full vaccine supply in NSW, Australia. *Vaccine* https://doi.org/10.1016/j.vaccine.2021.04.042 (2021).
51. Muller, C. P. Can integrated post-exposure vaccination against SARS-COV2 mitigate severe disease? *Lancet Reg. Health Euro.* **5**, 100118 (2021).
52. Althouse, B. M. et al. Superspreading events in the transmission dynamics of SARS-CoV-2: opportunities for interventions and control. *PLoS Biol.* **18**, e3000897 (2020).
53. Leclerc, Q. J. et al. What settings have been linked to SARS-CoV-2 transmission clusters? *Wellcome Open Res.* **5**, 83 (2020).

54. Christensen, H. et al. COVID-19 transmission in a university setting: a rapid review of modelling studies. *medRxiv* https://doi.org/10.1101/2020.09.07.20189688 (2020).
55. SPF. *Coronavirus: circulation des variants du SARS-CoV-2*. https://www.santepubliquefrance.fr/dossiers/coronavirus-covid-19/coronavirus-circulation-des-variants-du-sars-cov-2 (2021).
56. Metcalf, C. J. E. et al. Seven challenges in modeling vaccine preventable diseases. *Epidemics* **10**, 11–15 (2015).
57. Mistry, D. et al. Inferring high-resolution human mixing patterns for disease modeling. *Nat Commun* **12**, 323 (2021).
58. Béraud, G. et al. The French Connection: the first large population-based contact survey in France relevant for the spread of infectious diseases. *PLoS ONE* **10**, e0133203 (2015).
59. Google. *COVID-19 Community Mobility Report*. https://www.google.com/covid19/mobility?hl=en (2021).
60. YouGov. *Personal measures taken to avoid COVID-19*. https://yougov.co.uk/topics/international/articles-reports/2020/03/17/personal-measures-taken-avoid-covid-19 (2021).
61. Davies, N. G. et al. Age-dependent effects in the transmission and control of COVID-19 epidemics. *Nat. Med.* https://doi.org/10.1038/s41591-020-0962-9 (2020).
62. Di Domenico, L., Pullano, G., Sabbatini, C. E., Boëlle, P.-Y. & Colizza, V. Modelling safe protocols for reopening schools during the COVID-19 pandemic in France. *Nat. Commun.* **12**, 1073 (2021).
63. Riccardo, F. et al. Epidemiological characteristics of COVID-19 cases and estimates of the reproductive numbers 1 month into the epidemic, Italy, 28 January to 31 March 2020. *Euro. Surveill.* **25**, 2000790 (2020).
64. Kissler, S. M. et al. Viral dynamics of SARS-CoV-2 variants in vaccinated and unvaccinated persons. *N. Engl. J. Med.* **385**, 2489–2491 (2021).
65. Blanquart, F. et al. Characterisation of vaccine breakthrough infections of SARS-CoV-2 Delta and Alpha variants and within-host viral load dynamics in the community, France, June to July 2021. *Euro. Surveill.* **26**, 2100824 (2021).
66. Dagan, N. et al. BNT162b2 mRNA Covid-19 vaccine in a nationwide mass vaccination setting. *N. Engl. J. Med.* https://doi.org/10.1056/NEJMoa2101765 (2021).
67. Smith, L. E. et al. Adherence to the test, trace and isolate system: results from a time series of 21 nationally representative surveys in the UK (the COVID-19 Rapid Survey of Adherence to Interventions and Responses [CORSAIR] study). *medRxiv* https://doi.org/10.1101/2020.09.15.20191957 (2020).
68. Faucher, B. et al. *Code for Agent-based modelling of reactive vaccination of workplaces and schools against COVID-19*. https://doi.org/10.5281/zenodo.5910314 (Zenodo, 2022).

Acknowledgements

We acknowledge financial support from Haute Autorité de Santé; the ANR and Fondation de France through the project NoCOV (00105995); the Municipality of Paris (https://www.paris.fr/) through the programme Emergence(s); EU H2020 grants MOOD (H2020-874850; paper MOOD 035) and RECOVER (H2020-101003589) (the contents of this publication do not necessarily reflect the views of the European Commission); the ANRS through the project EMERGEN (ANRS0151); and the Institut des Sciences du Calcul et de la Donnée.

Author contributions

V.C., P.Y.B. and C.P. designed the analysis. P.Y.B. and C.P. developed the main methodology. B.F., R.A. and J.R. performed the analysis. D.L.B., C.T.K., S.C. and L.Z. critically commented on assumptions and model structure. P.Y.B. and C.P. drafted the manuscript. All authors discussed the results, edited and approved the contents of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-022-29015-y.

Correspondence and requests for materials should be addressed to Chiara Poletto.

Peer review information *Nature Communications* thanks Cameron Zachreson and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Supplementary Information

Agent-based modelling of reactive vaccination of workplaces and schools against COVID-19

Benjamin Faucher[1], Rania Assab[1], Jonathan Roux[2], Daniel Levy-Bruhl[3], Cécile Tran Kiem[4,5], Simon Cauchemez[4], Laura Zanetti[6], Vittoria Colizza[1,7], Pierre-Yves Boëlle[1], Chiara Poletto[1]

[1]Sorbonne Université, INSERM, Institut Pierre Louis d'Epidémiologie et de Santé Publique, F75012, France

[2]Univ Rennes, EHESP, CNRS, ARENES – UMR 6051, F-35000 Rennes, France

[3]Santé Publique France, Saint Maurice, France

[4]Mathematical Modelling of Infectious Diseases Unit, Institut Pasteur, Université de Paris, UMR2000, CNRS, Paris, France

[5]Collège Doctoral, Sorbonne Université, Paris, France

[6]Haute Autorité de Santé, Saint-Denis, France

[7]Tokyo Tech World Research Hub Initiative (WRHI), Tokyo Institute of Technology, Tokyo, Japan

Corresponding author: Chiara Poletto (chiara.poletto@inserm.fr)

Supplementary Methods

1. COVID-19 transmission model

We provide here in the following the parameter values for transmission and infection natural history (Supplementary Table 1), and the effect of vaccination (Supplementary Table 2). For a detailed explanation of the transmission model without vaccination we refer to[1]. Incubation period IP and length of the pre-symptomatic phase μ_p are specific for the Delta variant. In particular, μ_p is parametrised based on the proportion of pre-symptomatic transmission estimated in [2].

Supplementary Table 1. Transmission parameters and their baseline values.

Parameter	Description	Baseline value (other explored values)	Source
IP	Incubation period	5.8 days (5.1 days, 6.3 days)	[2] ([3,4])
$(\mu_p)^{-1}$	Average duration of the pre-symptomatic stage	2.1 days	Computed to recover 74% of pre-symptomatic transmission as estimated in [2]
$(\epsilon)^{-1}$	Rate of becoming infectious for exposed individuals	3.7 days	$IP - \mu_p^{-1}$
μ	Recovery rate	$(7\ \text{days})^{-1}$	[5]
β_I	transmissibility rescaling according to the infectious stage	0.51 for $I_{p,sc}$, I_{sc} 1 for $I_{p,c}$, I_c	[6]
ω_s	Transmission risk by setting	1 for household 0.3 for community 0.5 otherwise	[1]

Supplementary Table 2. Vaccine effectiveness parameters and their baseline values.

Parameter	Description	Baseline value (other explored values)	Source
$r_{S,1}$	reduction in susceptibility in the partial-protection stage	0.52 (0.35, 0.7)	from $VE_{S,1} = 48\%$, in the middle of the range of estimates after one dose in [7] (from 30% and 65%, worst and best estimates from [7], respectively)
τ_0	Average duration of the no-protection stage after first-dose inoculation	2 weeks (1 week)	
$r_{S,2}$	reduction in susceptibility in the maximum-protection stage	0.3 (0.2, 0.47)	from $VE_{S,2} = 70\%$, in the middle of the range of estimates after two doses in [7] (from 53% and 80%, worst and best estimates from [7], respectively)
$r_{c,2}$	reduction in the probability of developing clinical symptoms in the maximum-protection stage	0.9 (0.4, 0.95)	from $VE_{SP,2} = 73\%$, in the middle of the range of estimates after two doses in [7] (from 60% and 95%, worst and best estimates from [7], respectively)
τ_1	Average duration of the intermediate-protection stage	3 weeks (4 week, 8 week)	[8]
p_V	Probability of transition between $S^{V,1}$ and E^V	0.97 (0.65, 0.82)	Computed from the parameters above and assuming $VE_{SP,1} = 53\%$, in the middle of the range of estimates after one dose in [7] (from 35%, 75%, worst and best estimates from [7], respectively) (*)
r_I	Reduction in infection duration	25% (0)	[9]

(*) Mid range estimate in [7], $VE_{SP} = 55\%$, leads to a value of $p_V > 1$, thus $VE_{SP} = 53\%$ is taken.

2. Test-trace-isolate

We model case detection and isolation, combined with tracing and isolation of contacts according to the following rules:

- As an individual shows clinical symptoms, s/he is detected with probability $p_{d,c}$. If detected, case confirmation and isolation occur with rate r_d upon symptoms onset.
- Subclinical individuals are also detected with probability $p_{d,sc}$, and rate r_d .
- The index case's household contacts are isolated, with probability $p_{ct,HH}$, the same time the index case is detected and isolated. We assume that these contacts are tested at the time of isolation and among those all subclinical, clinical, pre-subclinical, and pre-clinical cases are detected (testing sensitivity 100%).
- Once the index case is detected, contacts of the index case occurring outside the household are traced and isolated with an average delay r_{ct}^{-1}. We define an acquaintance as a contact occurring frequently, i.e. with a frequency of activation higher than f_a. We assume that an acquaintance is detected and isolated with a probability $p_{ct,A}$, while other contacts (i.e. sporadic contacts) are detected and isolated with probability $p_{ct,sp}$, with $p_{ct,A} > p_{ct,sp}$. We assume that traced contacts are tested at the time of isolation and among those all subclinical, clinical, pre-subclinical, and pre-clinical cases are detected (testing sensitivity 100%).
- Only contacts (among contacts occurring both in household and outside) occurring within a window of D days before index case detection are considered for contact tracing.
- The index-case and the contacts are isolated for a duration d_I (for all infected compartments) and d_{NI} (for susceptible and recovered compartments). Contacts with no clinical symptoms have a daily probability p_{drop} to drop out from isolation.
- For both the case and the contacts, isolation is implemented by assuming no contacts outside the household and transmission risk per contact within a household reduced by a factor ι.

Parameter values are reported in Supplementary Table 3 and Supplementary Table 4 for baseline and enhanced TTI, respectively.

Supplementary Table 3. Model for test, trace, isolation. Parameters and their values for the baseline case.

Parameter	Description	Value	Source
$p_{d,c}$	Probability that a clinical case is detected	0.5	
$p_{d,sc}$	Probability that a subclinical case is detected	0.1	
r_d	For detected cases, rate of detection, confirmation and beginning of isolation	0.28 = (3.6 days)$^{-1}$	Average time from onset to testing is 2.6 days [10]. We assume one day to have the test results.
D	Length of the period preceding index-case confirmation, used to define a contact	6 days	$\simeq$2 days + ${r_d}^{-1}$ (a person is considered to be contact if s/he entered in contact with the index case during a window of 2 days preceding symptoms onset)
$p_{ct,HH}$	Probability that household contacts of an index case are identified and decide to isolate	0.7	
$p_{ct,A}$	Probability that acquittances of an index case are identified and decide to isolate	0.08	Calibrated to get $\simeq$2.8 contacts per index case on average (assumed $p_{ct,Oth} = p_{ct,A}/10$) [10]
$p_{ct,Oth}$	Probability that sporadic contacts of an index case are identified and decide to isolate	0.008	Calibrated to get $\simeq$2.8 contacts per index case on average (assumed $p_{ct,Oth} = p_{ct,A}/10$) [10]
r_{ct}	Rate of detection and isolation of contacts outside household	0.9 = (1.1 days)$^{-1}$	
f_a	Threshold frequency to define a contact as an acquaintance	1/7 days	
p_{do}	Probability to drop out from isolation for individuals that are not clinical	0.13	[11]
ι	Reduction in household transmission during isolation	0.5	

d_{NI}, d_I	Duration of isolation for an individual that is not infectious, duration of isolation for an individual that is infectious	10 days, 10 days	[12]

Supplementary Table 4. Model for test, trace, isolation. Parameters and their values for the case of enhanced TTI. Only values different from the baseline case are reported.

Parameter	Description	Baseline value
$p_{d,c}$	Probability that a clinical case is detected	0.7
$p_{d,sc}$	Probability that a subclinical case is detected	0.3
r_d	For detected cases, rate of detection, confirmation and beginning of isolation	0.9 = (1.1 days)$^{-1}$
$p_{ct,HH}$	Probability that household contacts of an index case are identified and decide to isolate	1
$p_{ct,A}$	Probability that acquittances of an index case are identified and decide to isolate	0.24
$p_{ct,Oth}$	Probability that sporadic contacts of an index case are identified and decide to isolate	0.024
p_{do}	Probability to drop out from isolation for individuals that are not clinical	0

3. Vaccination strategies

We provide here in the following the parameters values for the vaccination strategies detailed in the Methods section of the main paper.

Supplementary Table 5. Vaccine administering. Parameters and their baseline values.

Parameter	Description	Baseline values (other explored values)
$a_{th,V}$	Minimum age for vaccination	12 years
P_V^a	Probability an individual is willing to vaccinate	90% for 65+ 80% for <65 (60%, 100%)
n_{cl}	For reactive vaccination, cluster size for triggering reactive vaccination in a workplace or school	1 (2, 3, 4, 5)
T_{cl}	For reactive vaccination, time window for cluster definition	7 days
$(r_V)^{-1}$	For reactive vaccination, delay of implementation of the vaccination campaign once the cluster is identified (for workplaces/schools) and a case is identified (for households)	2 days (1 day, 4 days)
V_{TOT}	Maximum number of people vaccinated during one simulation	Unlimited
V_{day}	Maximum number of people vaccinated each day	Explored in the range [50, 500] per 100,000 inhabitants per day for mass, workplaces/universities, school locations Unlimited for reactive (explored in [50, 250] per 100,000 inhabitants per day)
$size_{th}$	For workplaces/universities, minimum size of a workplace that implement vaccination	20
n_{RV}	For reactive vaccination, cluster size for starting the reactive vaccination campaign in the flare up scenario	1, 5, 10
	Initial vaccination coverage	90% for 65+ 40% for <65 (15%, 25%, 35%, 45%, 55%, 65%, 75%)

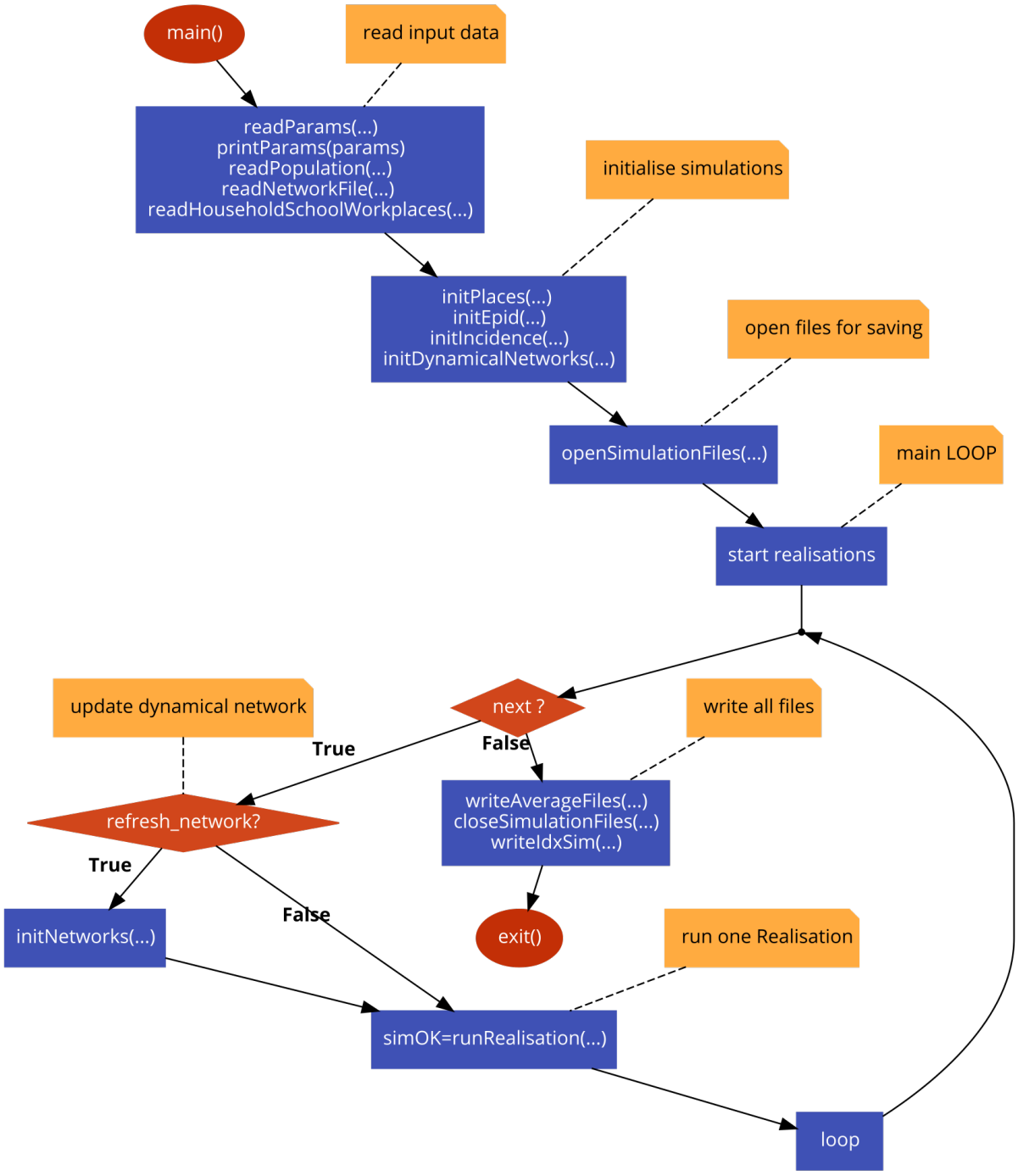
4. Details on the epidemic simulations

A schematic representation of the main program and of the simulation code and of the algorithm used for a single stochastic realisation are shown in Supplementary Figure 1 and 2, respectively. Simulations are discrete-time and stochastic. At each time step, corresponding to one day, three processes occur (Supplementary Figure 2):

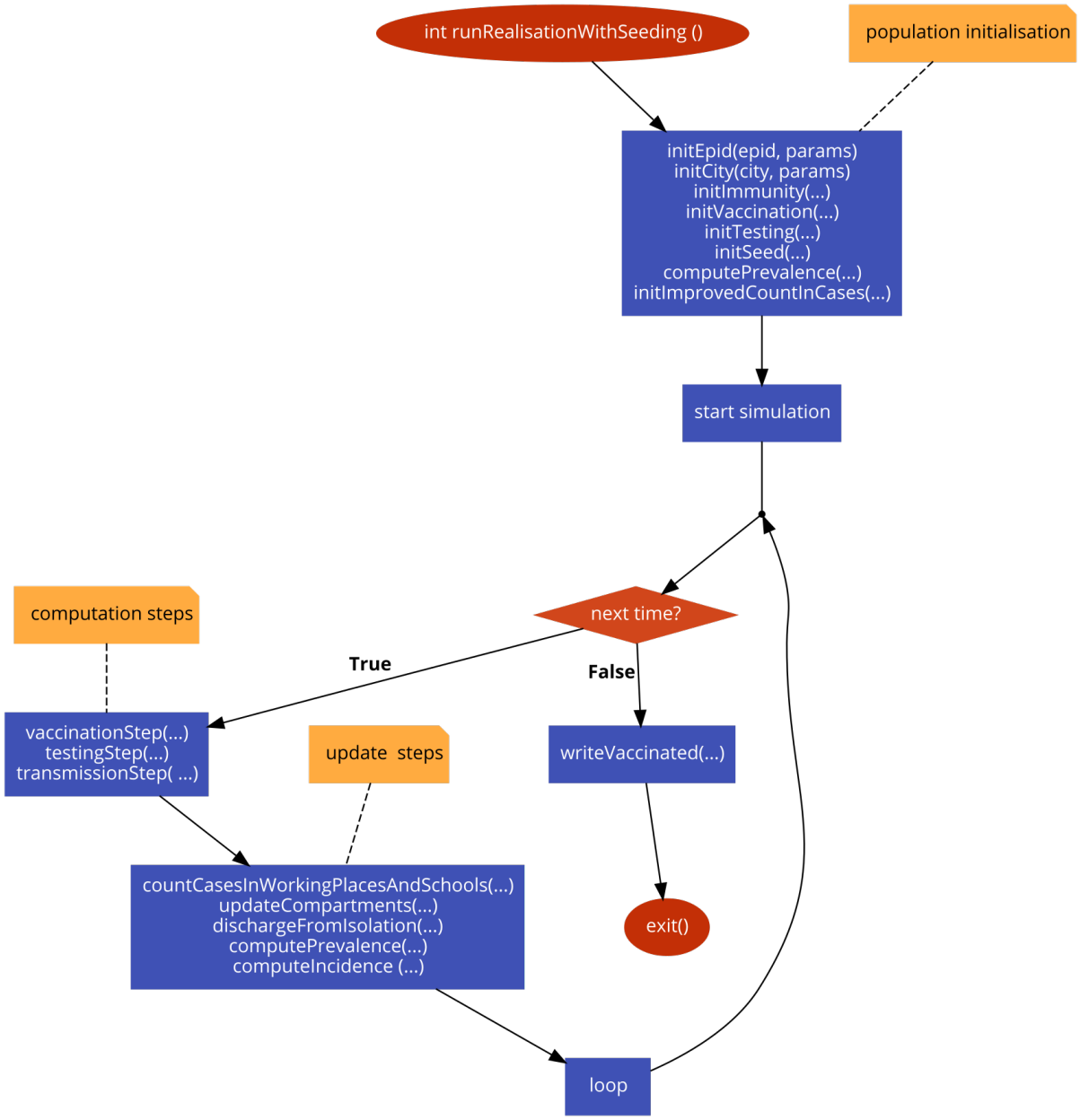
Vaccination Step: vaccines are administered according to the strategy or the strategies' combination.

Testing Step: cases are detected and isolated; contacts (within and outside household) are identified and isolated; isolated individuals get out from isolation.

Transmission Step: infectious status of nodes is updated. This includes transmission, recovery and transition through the different stages of the infection (e.g. from exposed to pre-symptomatic, from pre-symptomatic to symptomatic).



Supplementary Figure 1: Algorithm of the main program. Algorithm of the main program drawn with code2flow.com.



Supplementary Figure 2: Algorithm for one stochastic realisation. Algorithm for one stochastic realisation drawn with code2flow.com.

A single-run simulation is executed with no modelled intervention, until the desired immunity level is reached. This guarantees that immune individuals are realistically clustered on the network. We added some noise, by reshuffling the immune/susceptible status of 30% of the nodes to account for travelling, infection reintroduction from other locations and large gathering with consequent super-spreading not accounted for by the model. In Fig. 2 and 3 of the main

paper, all processes (transmission, TTI, vaccination) are simulated from the beginning of the simulation. In the flare-up scenario (Fig. 4), TTI and mass vaccination are modelled from the beginning. TTI is enhanced from the detection of the first case. Reactive vaccination starts after the detection of the first n_{RV}cases, with $n_{RV} = 1,5,10$ explored.

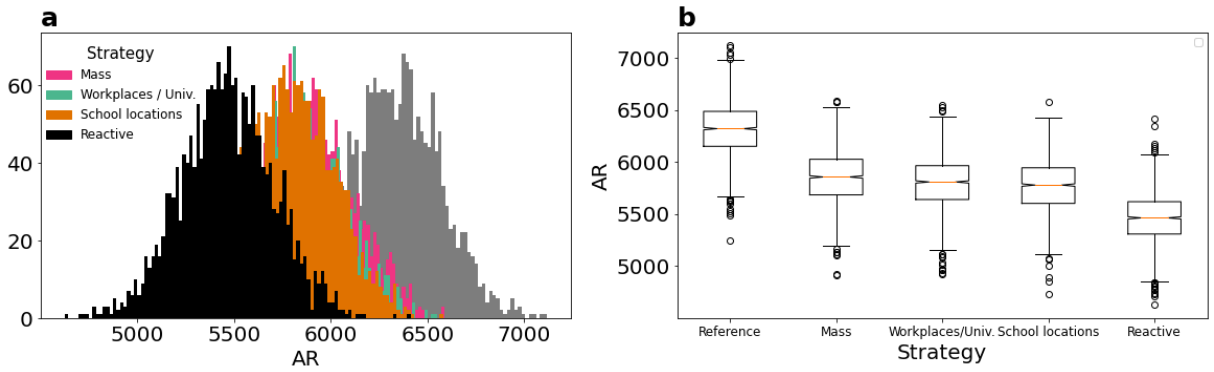
We vary COVID-19 transmission potential by tuning the daily transmission rate per contact β. The reproductive number R is computed numerically as the average number of infections each infected individual generates throughout its infectious period at the beginning of the simulation. Therefore, it integrates the effect of the interventions and the level of disease and vaccine induced immunity in the population at the start. We tune β to have the desired R value for the reference scenario, i.e. with only vaccination at the start. We then compare different vaccination strategies at the same value of transmissibility β.

To calibrate the duration of the pre-symptomatic stage from [2] (value reported in Supplementary Table 1) we generated an output file containing the list of all transmission events with the infection status of the infector. The proportion of pre-symptomatic transmission was computed as the fraction of transmission events with infector in the compartment $I_{p,sc}$ or $I_{p,c}$ over all infection events.

Supplementary Note 1: Comparison between reactive and non-reactive vaccination strategies

1. Distribution of the attack rates

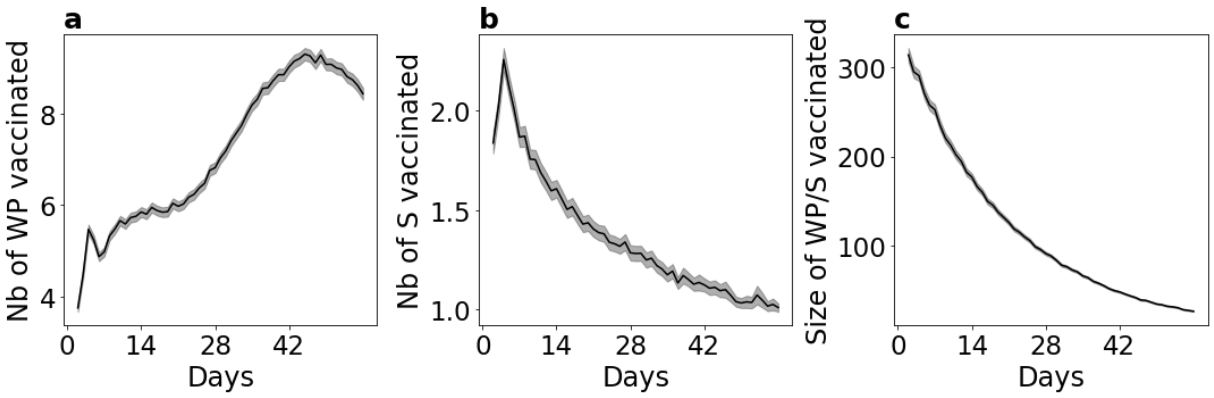
We compare here the distributions of the attack rates after two months across different strategies. In the Supplementary Figure 3, we consider the scenarios of Fig. 2e of the main paper and we show that the distribution and the standard deviation are similar among vaccination strategies.



Supplementary Figure 3: Distribution of the attack rate with different vaccination strategies. a Distribution of attack rates after two months for the scenarios analysed in Fig. 2e of the main paper. **b** Boxplots comparing the distributions of attack rates after 2 months for different vaccination strategies. Medians are represented by medium red lines and interquartiles (Q1-Q3) by boxes. The whiskers delimit the range between Q1 - 1.5*IQR and Q3 + 1.5*IQR. Outside values are considered as outcomes and are represented by points.

2. Vaccinated settings

We provide here a detailed analysis of the time evolution of the number of settings where vaccines are deployed in the context of reactive vaccination.

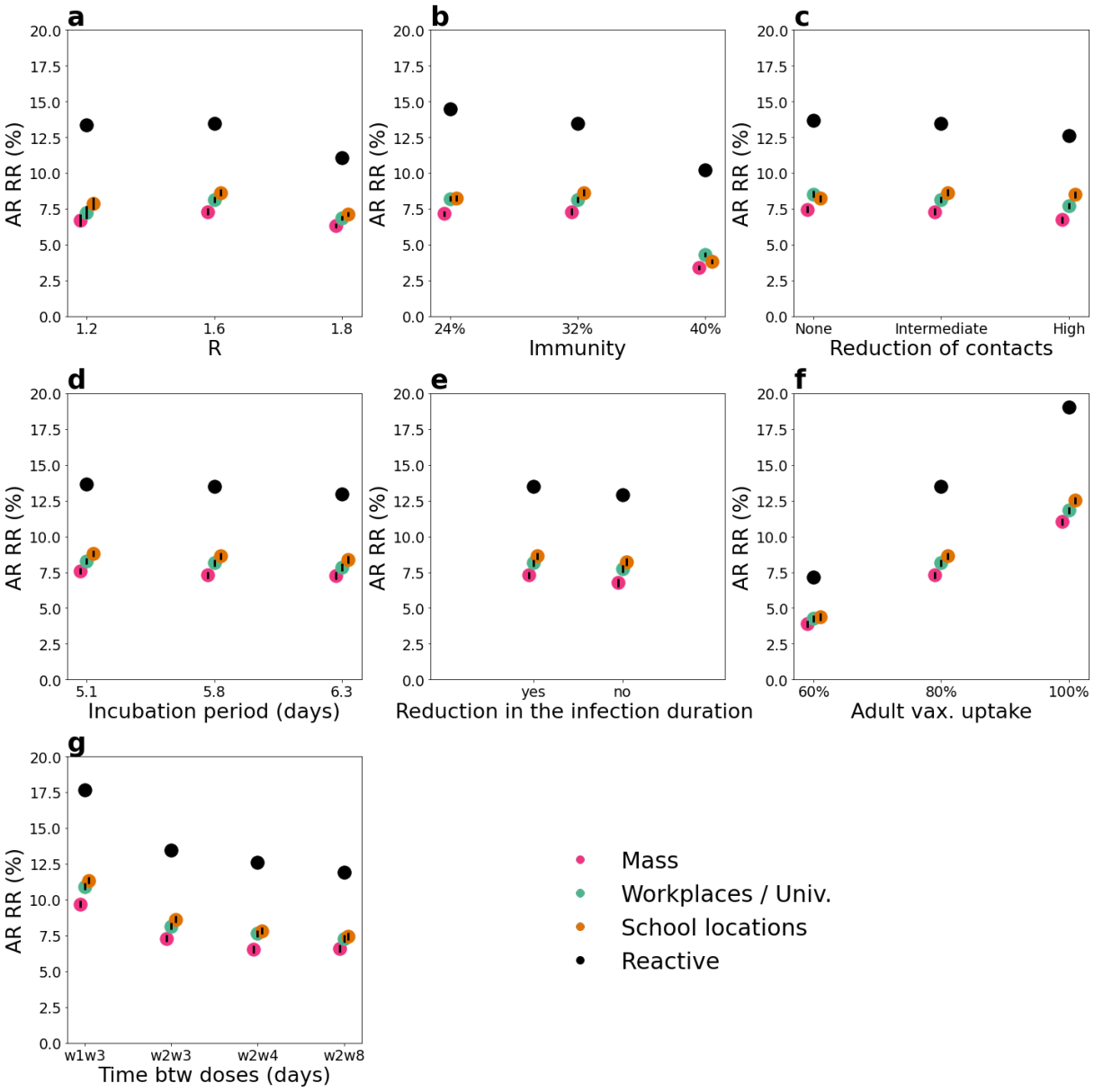


Supplementary Figure 4: Details analysis of settings where vaccines were deployed in the context of reactive vaccination. a, b Daily number of workplaces (**a**) and schools (**b**) where vaccines are deployed in the context of reactive vaccination. **c** Size of workplaces/Schools where vaccines are deployed as a function of time. Parameters are the same as in Fig. 2c of the main paper. In all panels continuous lines are means over 2000 independent stochastic realisations and the shaded areas are the standard error of the mean (+/- 2SEM).

3. Sensitivity analysis

We compare here reactive vaccination with non-reactive vaccination strategies under a variety of epidemic scenarios.

In Supplementary Figure 5 we compare all strategies at equal number of doses over the two-months period, exploring the impact of the following parameters: reproductive ratio, immunity level of the population, repartition of contacts across settings due to social distancing, incubation period, effect of the vaccine in the infection duration, vaccine uptake, and time between doses. Except when otherwise indicated, parameters are the ones of the baseline scenario with intermediate vaccination coverage (~45% of the whole population vaccinated, Fig. 2 d, e of the main paper). Increasing the transmissibility or initial immunity reduces the impact of reactive vaccination (panel a, b). In panel c we explore the impact of teleworking and reduction in community contacts by comparing the baseline scenario with scenarios with no or stronger restrictions. Reactive vaccination becomes more effective when no restriction is in place - although the effect is not strong. This is likely due to the enhanced role of workplaces as a setting of transmission when no teleworking is in place, thus bringing to an increased benefit of reactive vaccination targeting this setting. In panel d, we analyse the impact of the choice of the incubation period exploring values ranging from 5.1 [3] up to 6.3 [4]. We find that results are highly robust to the choice of the parameters within this range. In panel e we compare different hypotheses regarding the effect of the vaccine on the infection duration, i.e. the baseline case with the case in which the vaccine induces no reduction in the infectious duration. Also in this case, results are robust. Eventually in panel f we compare different levels of vaccine uptake - assuming uptake to be the same in the reactive and non-reactive strategies as in Fig. 2 of the main paper. The impact of vaccination increases with the uptake, the effect being stronger for the reactive strategy. Eventually, we compare in panel g different timing for vaccine-induced immunity to become effective. Specifically, we consider a case in which partial protection against infection mounts one week after the first dose. Assuming an incubation period of ~6 days, this would be consistent with no reduction in cases observed ~2 weeks after first-dose inoculation, as reported by some real vaccine effectiveness studies [13,14]. We then consider a longer interval between doses (i.e. 4 and 8 weeks). Assuming that protection against infection starts one week after first-dose inoculation leads to a higher impact of vaccination for all four vaccination strategies.

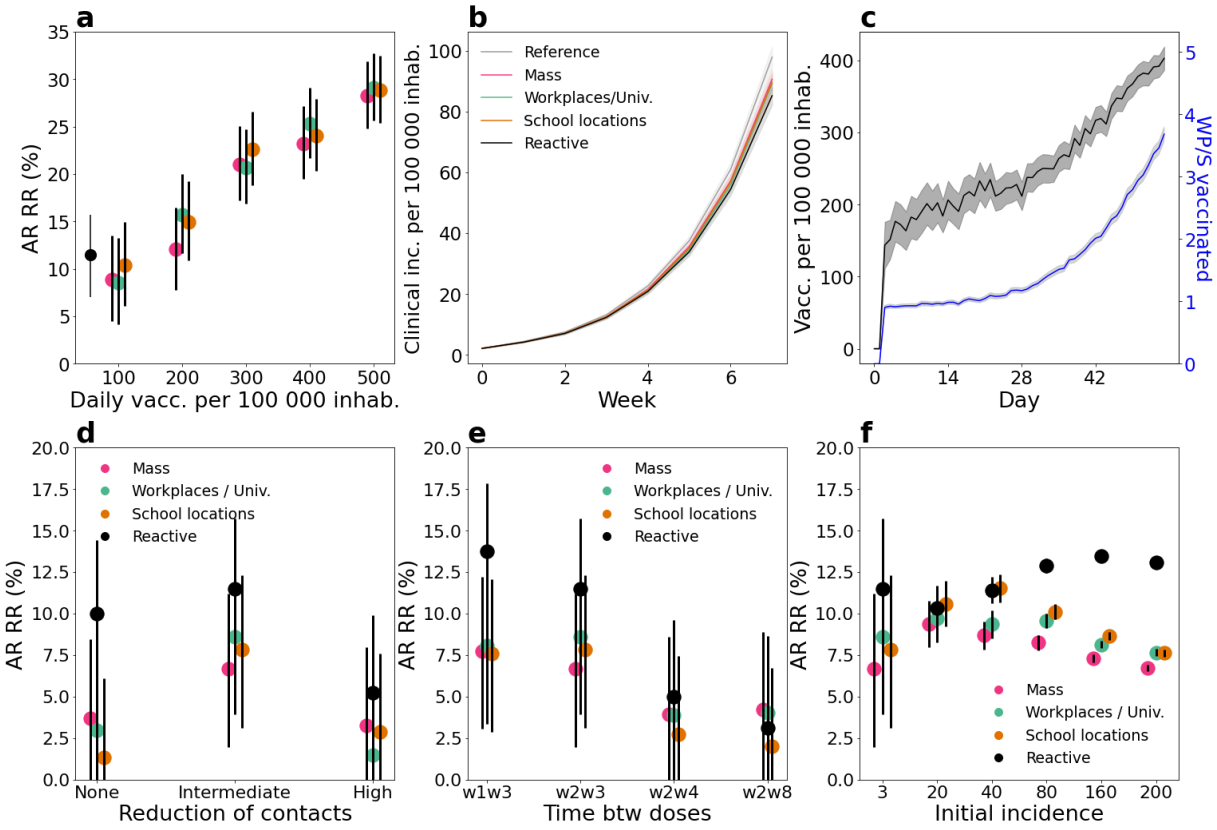


Supplementary Figure 5: Comparison between vaccination strategies - sensitivity analysis. Relative reduction (RR) in the attack rate (AR) after two months for all strategies at equal number of doses. RR is computed with respect to the reference scenario with initial vaccination only. Different parameter values and modelling assumptions are compared. Vaccination rate for mass, workplaces/universities and school location vaccination is set to the average value recovered for reactive vaccination. Exceptions made for the parameter explored in each panel - indicated in the x-axis -, all parameters are as in Fig. 2e. Parameters explored are: **a** reproductive ratio; **b** natural immunity of the population at the start; **c** repartition of contacts across settings due to social distancing (*Intermediate* is the baseline scenario, while *high* is given by a 30% reduction in community contacts and a 20% of individuals doing teleworking); **d** incubation period, **e** vaccine-induced reduction in infection duration (yes, no), **f** vaccine uptake, and **g** time between doses - the x-axis labels wNwM indicates the average number of weeks, N and M, of no protection following

first dose inoculation and of intermediate vaccine effectiveness, respectively. In all panels data are represented as means over 2000 independent stochastic realisations and error bars are derived from the standard error of the mean. These are smaller than the size of the dots in almost all cases.

The impact of reactive vaccination and its demand in terms of vaccine doses varies depending on the incidence level. In Supplementary Figure 6a-c we compare all strategies under a scenario of flare-up of cases. All parameters are as in the baseline scenarios, intermediate vaccination coverage (~45% of the whole population vaccinated, Fig. 2 d, e of the main paper), except for initial incidence. Here we assumed three cases were infected at the beginning. Panel a shows the relative reduction in the attack rate after two months as a function of the number of first daily doses, while panel b compares the incidence profiles under different strategies at equal number of vaccine doses. The relative reduction produced by reactive vaccination is close to the one produced by mass, school location and workplaces/universities. Panel c shows the number of vaccines deployed each day for reactive vaccination and the number of workplaces/schools where vaccines are deployed. These are initially low and increase gradually with incidence.

We then explore the impact of the reactive vaccination in the flare-up case in varying the different parameters. Specifically, we compare all strategies at an equal number of doses, varying the level of social distancing (Supplementary Figure 6d) and the timing for the immunity to mount after the first-dose vaccination (Supplementary Figure 6e). For certain parameter values reactive vaccination produces a higher relative reduction in the attack rate compared with non-reactive strategies. This is the case for instance when the development of vaccine immunity is rapid, and when no social restrictions are in place. In other cases it produces comparable effect. This is the case for instance of long delays between doses. Finally, in Supplementary Figure 6f, we provide an overview of the impact of initial incidence. As initial incidence increases the advantage of the reactive vaccination compared with the non-reactive strategies increases.



Supplementary Figure 6: Comparison between vaccination strategies - flare-up scenario. a Relative reduction (RR) in the attack rate (AR) over the first two months for all strategies as a function of the vaccination pace. RR is computed with respect to the reference scenario, with initial vaccination only. **b** Incidence of clinical cases with different vaccination strategies during the first 8 weeks. **c** Number of daily first-dose vaccinations, and workplaces/schools (WP/S in the plot) where vaccines are deployed. **d** AR RR after two months according to the repartition of contacts across settings due to social distancing - *Intermediate* is the baseline scenario, while *high* is given by a 30% reduction in community contacts and a 20% of individuals doing teleworking. **e** AR RR after two months according to the timing for the immunity to mount after first-dose vaccination - the x-axis labels wNwM indicates the average number of weeks, N and M, of no protection following first dose inoculation and of intermediate vaccine effectiveness, respectively. **f** AR RR after two months according to the initial incidence for all strategies at equal number of doses. In panels **b,d-f** all strategies are compared at equal number of doses. All panels, except for panel **f**, consider a flare up scenario, where the epidemic is seeded with 3 infectious individuals. All other parameters are as in Fig. 2d, e of the main paper. In panels **a**, **d-f** data are represented as means over 2000 independent stochastic realisations and error bars are derived from the standard error of the mean. In panels **b**, **c** continuous lines are means over 2000 independent stochastic realisations and the shaded areas are the standard error of the mean (+/- 2SEM).

4. Additional epidemic outcomes

Based on the estimated incidence of clinical cases per day provided by the transmission model, we infer outcomes related to hospital, namely hospital and intensive care unit (ICU) entries, beds in ICU ward, and deaths. We use age-dependent hospital admissions (ICU and non-ICU) risks estimated by [15,16] and ICU admission risks for hospitalised patients based on SI-VIC extract [17]. Hospital admissions risks were adjusted to apply only to clinical cases [18] and to account for vaccine effectiveness for hospitalisation for zero, half (1 dose) and full (2 doses) vaccination. We assume the vaccine efficacies for hospitalisation were equal to 83% and 94% for half and full vaccinations, respectively[7]. We also assume that the hospital admission risk increases by 80% with the Delta variant compared to Alpha [19] and by 40% with the Alpha variant compared to the wild type [20]. Patients who were hospitalised entered the hospital on average 7 days (sd = 3.9 days – Gamma distribution) after the beginning of the infectious phase [21]. Those who were admitted in ICU enter this unit with a mean delay of 1.69 days (assuming an exponential distribution) [17]. To estimate the number of occupied beds, we use age-specific mean durations of stay and their corresponding standard deviations in ICU calculated on all the hospitalised cases in the first 9 months of the French epidemic (March-November 2020)[17]. We assume that the standard deviations of ICU lengths of stay were equal to the corresponding mean and do not consider post-ICU care in the estimation of occupied beds. We estimate the number of deaths using hospital and ICU death risks of hospitalised infected persons [17]. Deaths are delayed in time using the mean delays and standard deviations from hospital or ICU admission to death [17]. All lengths of stay are supposed to follow a Gamma distribution. Parameters and their values are summarised in Supplementary Tables 6 and 7.

We also estimate the number of life years and quality-adjusted life years (QALY) lost for each death using life-tables provided by 'French National Institute of Statistics and Economic Studies' (INSEE) for 2012-2016 [22] and utility measures of each age-group in France [23].

Supplementary Table 6. Risks of hospitalisation according to vaccination status, ICU admission and death in general ward and ICU

Age group	Hospitalisation risk for no-vaccination	Hospitalisation risk for 1-dose vaccination	Hospitalisation risk for 2-doses vaccination	ICU admission risk	Death risk in general ward	Death risk in ICU
0-9	0.0246	0.0115	0.0066	0.159	0.006	0.078
10-19	0.0136	0.0063	0.0037	0.159	0.006	0.078
20-29	0.0364	0.0170	0.0098	0.159	0.006	0.078
30-39	0.0443	0.0207	0.0119	0.159	0.006	0.078
40-49	0.0617	0.0288	0.0166	0.219	0.016	0.103
50-59	0.1697	0.0792	0.0457	0.270	0.030	0.175
60-69	0.2360	0.1101	0.0635	0.299	0.064	0.268
70-79	0.5113	0.2386	0.1377	0.235	0.140	0.366
79+	0.9496	0.4431	0.2557	0.053	0.308	0.468

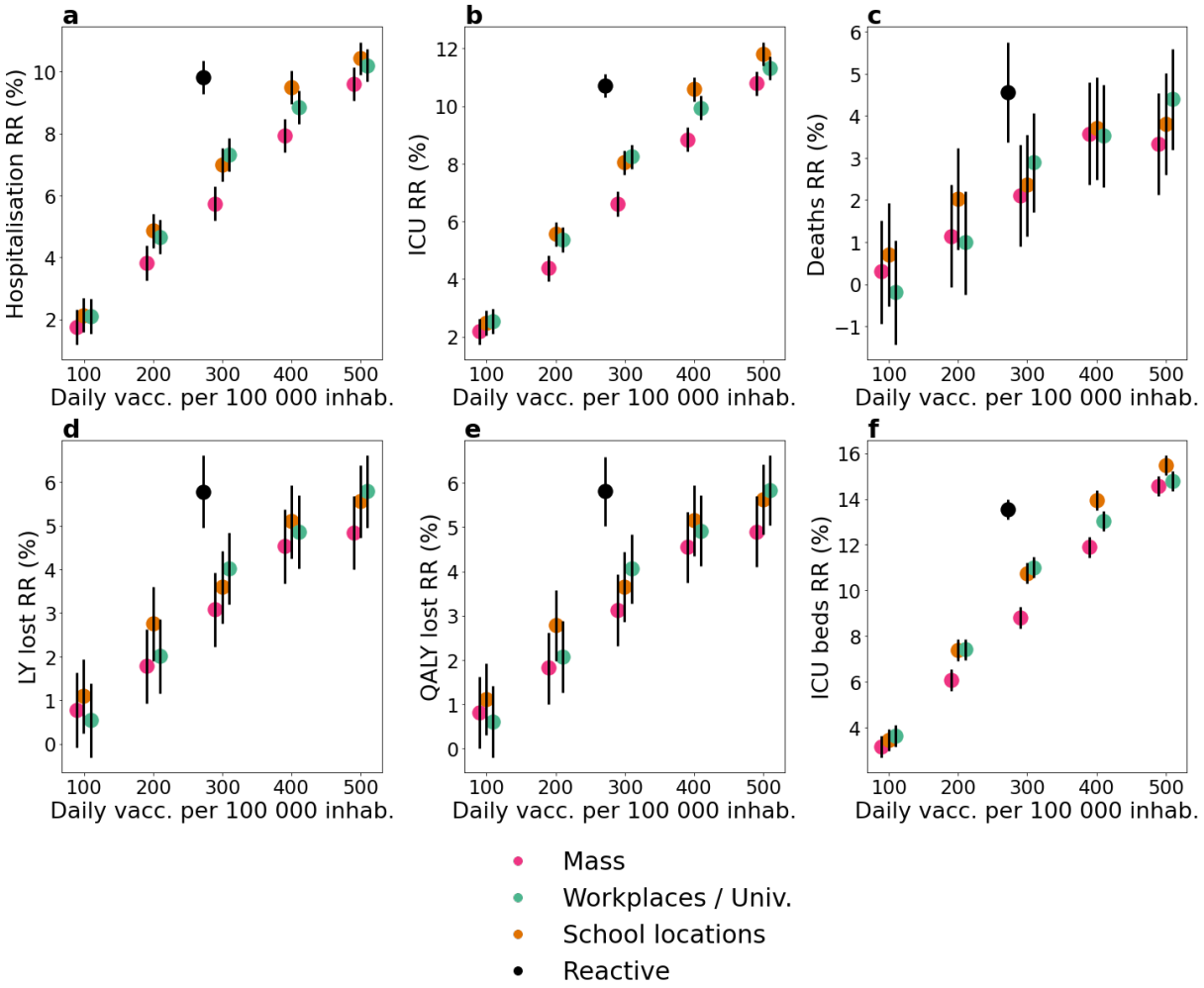
Supplementary Table 7. Delays in days from hospitalisation admission in general ward to death or hospital discharge and delays from ICU admission to ICU discharge or death.

Age group	Mean los in general ward (sd)	Mean los in general ward for dying people (sd)	Mean los in ICU (sd)	Mean los in ICU for dying people (sd)
0-9	6.4 (6.8)	10.6 (12.3)	12.7 (12.7)	22.3 (22.3)

10-19	6.4 (6.8)	10.6 (12.3)	12.7 (12.7)	22.3 (22.3)
20-29	6.4 (6.8)	10.6 (12.3)	12.7 (12.7)	22.3 (22.3)
30-39	6.4 (6.8)	10.6 (12.3)	12.7 (12.7)	22.3 (22.3)
40-49	6.4 (6.8)	10.6 (12.3)	12.7 (12.7)	22.3 (22.3)
50-59	6.4 (6.8)	10.6 (12.3)	12.7 (12.7)	22.3 (22.3)
60-69	9.3 (9.1)	10.6 (12.3)	16.7 (16.7)	20.8 (20.8)
70-79	11.7 (11.4)	10.6 (12.3)	17.5 (17.5)	18.9 (18.9)
79+	15 (13.8)	10.6 (12.3)	13.6 (13.6)	10.6 (10.6)

Los: Length of stay. Sd = Standard deviation.

Supplementary Figure 7 shows the relative reductions in the number of hospitalisations, deaths, ICU entries, life-year lost, quality-adjusted life-year lost and ICU bed occupancy at the peak, comparing each vaccination scenario with the reference scenario - i.e. vaccination only at the start. We consider here the high incidence and intermediate vaccine coverage scenario, analogously to Fig. 2d, e of the main paper. The six indicators show a behaviour similar to incidence. Overall reduction values are smaller. This is expected, since a large proportion of elderly are already vaccinated at the start, and the compared vaccination strategies target a population that is less at risk of severe infection. Still, all indicators show the same qualitative behaviour, with reactive vaccination outperforming the non-reactive vaccination strategies at equal number of first-dose vaccination.

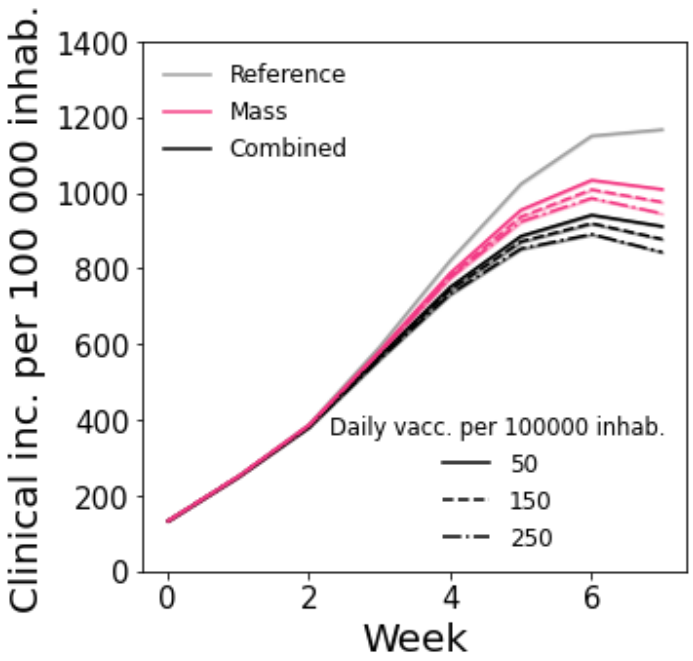


Supplementary Figure 7: Comparison between vaccination strategies - additional epidemic outcomes. Relative reduction (RR) in the cumulative incidence of: **a** hospitalisations, **b** intensive care unit (ICU) entries, **c** deaths, **d** life years (LY) lost and **e** quality-adjusted life years (QALY) lost over the first two months for all strategies as a function of the vaccination pace. **f** Relative reduction (RR) in occupied ICU beds at the peak over the first two months for all strategies as a function of the average daily number of first-dose vaccinations. We consider here the baseline scenario with intermediate vaccination coverage - i.e. same parameters as in Fig. 2d, e. In all panels, data are represented as means over 2000 independent stochastic realisations and error bars are derived from the standard error of the mean.

Supplementary Note 2: Combined reactive and mass vaccination for managing sustained COVID-19 spread

1. Additional results

In Supplementary Figure 8 we show the incidence curve corresponding to the scenarios analysed in Fig. 3a of the main paper. Mass and combined vaccination with three different vaccination paces are compared.

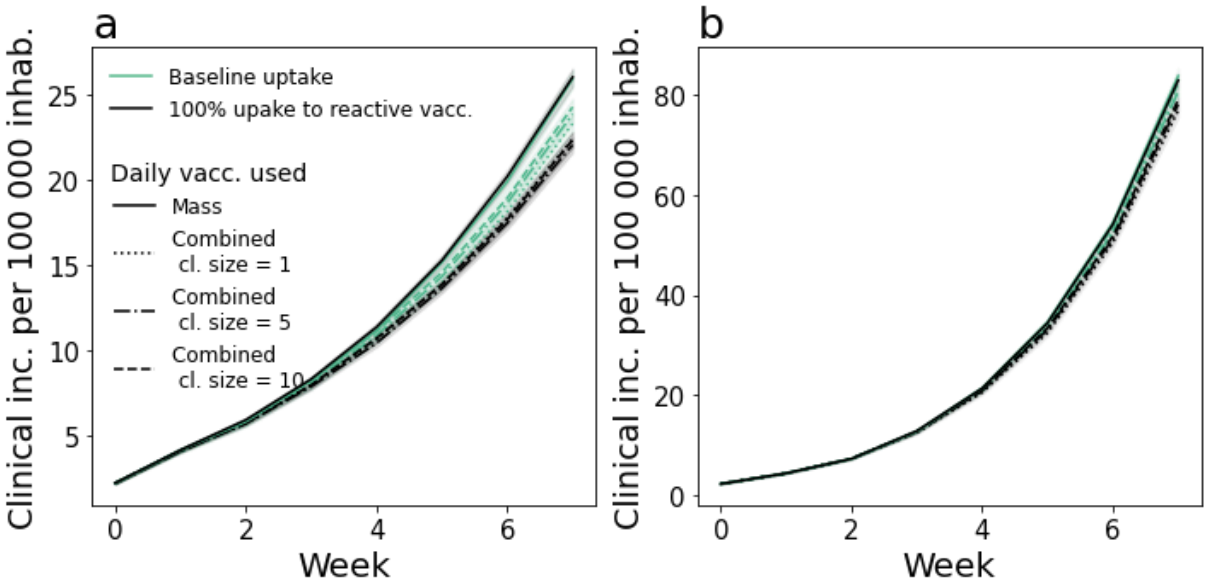


Supplementary Figure 8: Incidence of clinical cases for mass and combined vaccination strategies in the case of sustained spread. Scenarios are the same as the ones plotted in Fig. 3a of the main paper. Continuous lines are means over 2000 independent stochastic realisations and error bands are derived from the standard error of the mean (+/- 2SEM) - this is very low for this set of parameters.

Supplementary Note 3: Combined reactive and mass vaccination for managing a COVID-19 flare-up

1. Additional results

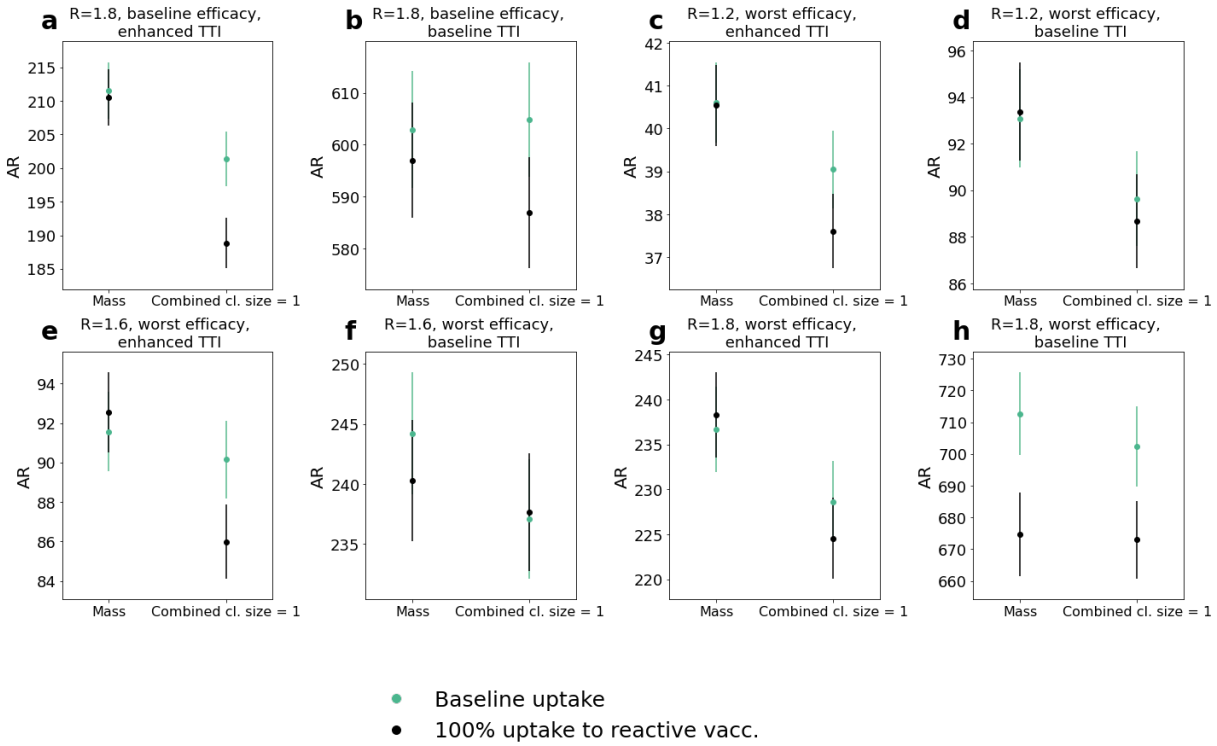
In Supplementary Figure 9 we show the incidence curve corresponding to the scenarios analysed in Fig. 4 of the main paper. Mass and combined vaccination with the different vaccination scenarios considered are compared.



Supplementary Figure 9: Incidence of clinical cases for mass and combined vaccination strategies in the case of flare-up. Scenarios are the same as the ones plotted in Fig. 4 of the main paper. **a** Scenario with enhanced TTI. **b** Scenario with baseline TTI. In both panels, continuous lines are means over 8000 independent stochastic realisations and error bands are derived from the standard error of the mean.

2. Sensitivity analysis

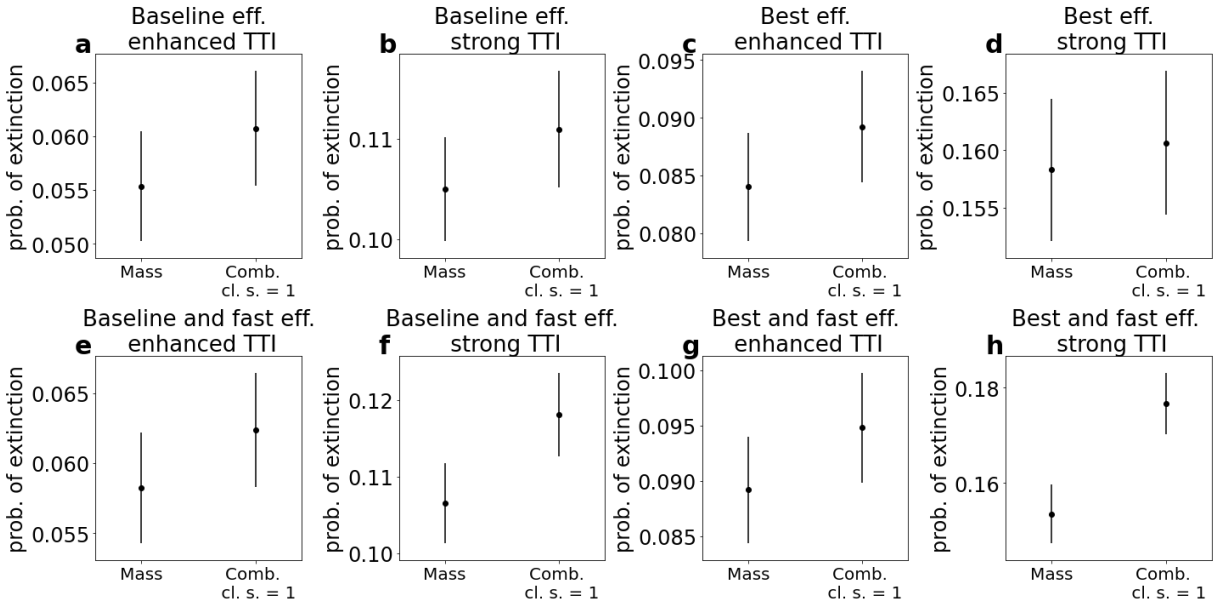
In Supplementary Figure 10 we analyse the impact of combined and mass vaccination in a flare up scenario similarly to Fig. 4 of the main paper, by varying the hypotheses on virus transmissibility and vaccine escape. Specifically we consider values of the reproductive ratio from 1.2 to 1.8, and both worst and baseline vaccine effectiveness level - the worst vaccine effectiveness level is the same as in Fig. 2i of the main paper. Analogously to Fig. 4 we compare the attack rate for combined and mass vaccination, assuming both baseline and enhanced TTI and both baseline and 100% vaccine uptake in the context of reactive vaccination. We consider only the case in which reactive vaccination starts after the detection of the first case. For each set of parameters, scenarios with enhanced TTI and 100% uptake are associated with smaller attack rates and larger difference between mass and combined than the corresponding scenarios with baseline TTI and baseline uptake.



Supplementary Figure 10: Combined reactive and mass vaccination for managing a COVID-19 flare-up - sensitivity analysis. Average attack rate per 100000 inhabitants after two months for the flare-up case under different hypotheses: **a, b,** $R = 1.8$ with baseline vaccine effectiveness; **c-h** worst vaccine effectiveness with $R = 1.2$ (**c,d**), $R = 1.6$ (**e,f**) and $R = 1.8$ (**g,h**). The worst vaccine effectiveness scenario is defined as in Fig. 2 i of the main paper, i.e. $VE_{S,1} = 30\%$, $VE_{SP,1} = 35\%$, $VE_{S,2} = 53\%$ and $VE_{SP,2} = 60\%$. We compare enhanced and baseline TTI - a, c, e, g and b, d, f, h, respectively -, as well as baseline and 100% vaccine uptake. In all panels, parameters are the same as in Fig. 4 except for otherwise indicated. In all panels, data are represented as means over 8000 independent stochastic realisations and error bars are derived from the standard error of the mean.

We then investigate the impact of combined and mass vaccination on the extinction of the flare-up. In Supplementary Figure 11a we plot the probability of extinction for the scenario considered in Fig. 4a of the main paper (enhanced TTI and 100% vaccine uptake). We find that the probability of extinction is ~5%, and the difference between mass and combined is ~0.5%. The probability of extinction progressively increases under more optimistic hypotheses: increase in case detection from 70% and 30% (enhanced TTI) to 100% and 50% (strong TTI) for clinical and subclinical cases, respectively; increase in vaccine effectiveness to the best case scenario considered in Fig. 2i; rapid mounting of the vaccine effect, with partial immunity against infection already present one week after inoculation. In the best-case scenario plotted in panel h, the

probability of extinction reaches ~0.15 and ~0.18 for mass and combined vaccination, respectively.



Supplementary Figure 11: Combined reactive and mass vaccination for managing a COVID-19 flare-up - probability of extinction. Three sets of parameters are investigated. 1) Baseline vaccine effectiveness (**a,b,e,f**) vs best effectiveness, i.e. $VE_{S,1} = 65\%$, $VE_{SP,1} = 75\%$,$VE_{S,2} = 80\%$ and $VE_{SP,2} = 95\%$, (**c,d,g,h);** 2) enhanced TTI (**a,c,e,g**) vs. strong TTI with $p_{d,c}$= 1 and $p_{d,sc}$= 0.5 (**b,d,f,h**). 3) 2 weeks (baseline) for vaccines to reach partial effectiveness (**a,b,c,d**) vs 1 week (**e,f,g,h**). In all panels, parameters are the same as in Fig. 4. In all panels, values represent the fraction of 15000 stochastic runs where the epidemic ends before two months. Error bars represent the standard error assuming the number of extinctions follows a binomial distribution.

Supplementary References

1. Moreno López, J. A. *et al.* Anatomy of digital contact tracing: Role of age, transmission setting, adoption and case detection. *Science Advances* eabd8750 (2021) doi:10.1126/sciadv.abd8750.
2. Kang, M. *et al. Transmission dynamics and epidemiological characteristics of Delta variant infections in China*. 2021.08.12.21261991 https://www.medrxiv.org/content/10.1101/2021.08.12.21261991v1 (2021) doi:10.1101/2021.08.12.21261991.

3. Lauer, S. A. *et al.* The Incubation Period of Coronavirus Disease 2019 (COVID-19) From Publicly Reported Confirmed Cases: Estimation and Application. *Ann Intern Med* **172**, 577–582 (2020).
4. Xin, H. *et al.* The Incubation Period Distribution of Coronavirus Disease 2019: A Systematic Review and Meta-analysis. *Clinical Infectious Diseases* (2021) doi:10.1093/cid/ciab501.
5. He, X. *et al.* Temporal dynamics in viral shedding and transmissibility of COVID-19. *Nature Medicine* 1–4 (2020) doi:10.1038/s41591-020-0869-5.
6. Di Domenico, L., Pullano, G., Sabbatini, C. E., Boëlle, P.-Y. & Colizza, V. Impact of lockdown on COVID-19 epidemic in Île-de-France and possible exit strategies. *BMC Medicine* **18**, 240 (2020).
7. Higdon, M. M. *et al. A systematic review of COVID-19 vaccine efficacy and effectiveness against SARS-CoV-2 infection and disease*. 2021.09.17.21263549 https://www.medrxiv.org/content/10.1101/2021.09.17.21263549v1 (2021) doi:10.1101/2021.09.17.21263549.
8. Polack, F. P. *et al.* Safety and Efficacy of the BNT162b2 mRNA Covid-19 Vaccine. *New England Journal of Medicine* **383**, 2603–2615 (2020).
9. Kissler, S. M. *et al.* Viral Dynamics of SARS-CoV-2 Variants in Vaccinated and Unvaccinated Persons. *New England Journal of Medicine* **385**, 2489–2491 (2021).
10. SPF. COVID-19 : point épidémiologique du 6 mai 2021. https://www.santepubliquefrance.fr/maladies-et-traumatismes/maladies-et-infections-respiratoires/infection-a-coronavirus/documents/bulletin-national/covid-19-point-epidemiologique-du-6-mai-2021.
11. Smith, L. E. *et al.* Adherence to the test, trace and isolate system: results from a time series of 21 nationally representative surveys in the UK (the COVID-19 Rapid Survey of Adherence to Interventions and Responses [CORSAIR] study). *medRxiv* (2020) doi:10.1101/2020.09.15.20191957.

12. Info Coronavirus COVID-19 - Tester - Alerter - Protéger. *Gouvernement.fr* https://www.gouvernement.fr/info-coronavirus/tests-et-depistage.

13. Nasreen, S. *et al. Effectiveness of mRNA and ChAdOx1 COVID-19 vaccines against symptomatic SARS-CoV-2 infection and severe outcomes with variants of concern in Ontario*. 2021.06.28.21259420 https://www.medrxiv.org/content/10.1101/2021.06.28.21259420v3 (2021) doi:10.1101/2021.06.28.21259420.

14. Tang, P. *et al.* BNT162b2 and mRNA-1273 COVID-19 vaccine effectiveness against the SARS-CoV-2 Delta variant in Qatar. *Nat Med* 1–8 (2021) doi:10.1038/s41591-021-01583-4.

15. Kiem, C. T. *et al.* Short and medium-term challenges for COVID-19 vaccination: from prioritisation to the relaxation of measures. (2021).

16. Lapidus, N. *et al.* Do not neglect SARS-CoV-2 hospitalization and fatality risks in the middle-aged adult population. *Infectious Diseases Now* **51**, 380–382 (2021).

17. Lefrancq, N. *et al.* Evolution of outcomes for patients hospitalised during the first 9 months of the SARS-CoV-2 pandemic in France: A retrospective national surveillance data analysis. *The Lancet Regional Health – Europe* **5**, (2021).

18. Davies, N. G. *et al.* Age-dependent effects in the transmission and control of COVID-19 epidemics. *Nature Medicine* 1–7 (2020) doi:10.1038/s41591-020-0962-9.

19. Sheikh, A., McMenamin, J., Taylor, B. & Robertson, C. SARS-CoV-2 Delta VOC in Scotland: demographics, risk of hospital admission, and vaccine effectiveness. *Lancet* **397**, 2461–2462 (2021).

20. Bager, P. and Wohlfahrt, J. and Rasmussen, M. and Albertsen, M and Grove Krause, T. Hospitalisation associated with SARS-CoV-2 delta variant in Denmark. *Lancet Infect Dis* (2021).

21. Salje, H. *et al.* Estimating the burden of SARS-CoV-2 in France. *Science* (2020) doi:10.1126/science.abc3517.

22. Tables de mortalité par sexe, âge et niveau de vie − Tables de mortalité par niveau de vie | Insee. https://www.insee.fr/fr/statistiques/3311422?sommaire=3311425.

23. Chevalier, J. & de Pouvourville, G. Valuing EQ-5D using time trade-off in France. *Eur J Health Econ* **14**, 57–66 (2013).

4.4 Conclusion.

In this chapter, we showed that in most scenarios, reactive vaccination outperformed non-reactive strategies in mitigating the epidemic for an equal number of vaccine doses. A combined approach, with both mass and reactive vaccination, was also more effective than mass vaccination only. We tested the impact of input parameters on the effectiveness of reactive vaccination to understand in what situation this vaccination strategy was indicated. Overall, parameters affecting the number of people vaccinated around a case have a higher impact on the effectiveness of reactive vaccination. For instance, if the number of vaccinated people is high already at the onset of the simulation, there will be fewer candidates for reactive vaccination around a detected case. In this case, reactive vaccination will be less advantageous than when fewer people are initially vaccinated. Conversely, changing the reproductive number, the natural immunity at the start or the reduction of contacts due to teleworking or limitation of social activities showed little impact on the relative advantage of reactive vaccination. Vaccine uptake had an important impact on the effectiveness of reactive vaccination when aiming at both mitigation and control. Vaccine hesitancy has proven to be a significant obstacle to vaccination policies, particularly in France [170, 171, 172]. Psychological studies have shown that individuals who perceive a low risk of illness are the most likely to be hesitant about vaccination. Therefore, we could expect that a reactive strategy could increase adherence to vaccination. Knowing the risk of infection within a school or workplace could encourage individuals attending those places to get vaccinated. Lower vaccination rate is also associated with other factors such as low socio-economic level, difficult access to vaccination or distrust in health authorities [173, 174]. In response, reactive vaccination could be part of an "Aller-vers" policy [175] that involves going out and vaccinating people who are less likely to go to vaccination centres themselves where they live.

Chapter 5

Conclusion and discussion.

The emergence of new pathogens poses important challenges to public health authorities. During the COVID-19 pandemic, with the appearance of the new coronavirus in China in 2019 and the emergence of new variants, prompt and appropriate reaction was needed to respond to the threat of new outbreaks. However, the heterogeneity of surveillance systems has made it difficult to interpret surveillance data and plan appropriate responses. In this thesis, we developed mathematical models to deal with this problem. In the work presented in 3, we presented a retrospective study aimed at better interpreting surveillance data for epidemic reconstruction. We took the emergence of the Alpha variant as a case study to understand the drivers and impacts of the silent spread. More precisely, we used GISAID metadata of Alpha sequences to reconstruct Alpha dissemination from the United Kingdom to other countries and estimated the importation date from the first detection date. We explored the interaction between sequencing coverage, air travel volumes, and delay from collection to submission, in determining the duration of silent circulation. By integrating these factors, we reconstructed the spread of the Alpha variant, and we showed that it circulated silently for days to months and reached more than 60 countries by the end of 2020 to be compared with the 24 that reported cases. Finally, we used two different models to reproduce the local spread of the Alpha variant in six countries [8] [9, 10, 11]. We showed that countries implementing mitigation measures likely delayed the establishment of local transmission, mitigating the negative consequences of late detection. While this study was retrospective, our framework could be used in future pandemics to monitor the unfolding propagation of a new virus. Our results suggest that once a new pathogen emerges, other countries should not wait to detect the virus before implementing measures. In particular, enhancing screening and implementing local mitigation measures could delay the variant propagation and decrease the impact of silent spread.

Various mitigation measures were put in place during the pandemic: limitation or ban of public gatherings, curfews, closure of certain shops, closure of schools, target vaccination or strict lockdowns. However, these measures had important economic and social costs [176, 177]. Uncertainty about the balance between the benefits and costs of a mitigation measure can delay the implementation resulting in avoidable cases and deaths. Mathematical models can help to propose scenarios of intervention and quantify costs and benefits. In the second work of this thesis, we proposed a model to assess the effectiveness of reactive vaccination in workplaces and schools. We used an agent-based model, modelling the daily interaction of individuals in different settings. We accounted for various factors: individual immunity, contact network structure, number of available doses, and concomitant non-pharmaceutical interventions. In most scenarios, reactive vaccination proved to be more effective than mass or targeted vaccination in reducing the incidence of clinical cases over two months. In particular, the comparative advantage was more pronounced in situations with low

initial vaccination coverage or high vaccination uptake. We also examined a combined strategy, where reactive vaccination is implemented on the top of mass vaccination policy. Once again, the strategy involving reactive vaccination outperformed mass vaccination alone in most scenarios. We also computed the number of daily vaccine doses used. We found that on average over the period the reactive strategy saved doses compared with mass vaccination. In case of a flare-up, like the emergence of a new VOC, the combined strategy could help to limit the spread only if the vaccination campaign is supported by a strong Test-Trace-Isolate system and an increase in vaccination uptake. Provided that sufficient data are available, our model could be adapted to evaluate reactive vaccination or targeted vaccination for any epidemics caused by respiratory viruses like influenza. We could also evaluate the impact of contact tracing policies, and test the impact of the detection rate, the delay in detection, or the probability to trace contacts.

The work carried out in this thesis was conducted in a specific context during the COVID-19 pandemic. The two studies, beyond their theoretical interest, aimed to provide information relevant for public health authorities regarding concrete situations. The research on the Alpha variant aimed to understand the impact of available data on monitoring the spread of a new virus. The second study sought to determine whether the strategy of reactive vaccination could be implemented in response to the emergence of the Delta variant in France. This work had to be adapted to the real-time epidemic situation. The model was initially calibrated in early 2021 on the Alpha variant. We had to recalibrate the model parameters for the arrival of the Delta variant in the summer of 2021 , and adapt certain results. We also took into account variations in contact due to responses and measures to face waves of cases.

We used a highly detailed model to closely align with the situation being studied. Numerous parameters were examined to understand the feasibility of implementing reactive vaccination. In addition to the effectiveness of this strategy in reducing the attack rate, we tested the influence of various logistical parameters: detection rate, number of cases required to trigger the strategy or the time required to implement vaccination. Overall, we demonstrated that in the studied case, reactive vaccination could be beneficial for mitigating the epidemic but would not be effective in containing the spread of a new variant. Logistical parameters that do not impact the number of people to be vaccinated have little effect on the effectiveness of reactive vaccination. However, the decision to implement this strategy requires balancing these conclusions with data on available resources in practice.

The two works presented in this thesis relied on pre-existing frameworks but were adapted to the COVID-19 situation by integrating data collected during the pandemic. The framework used to study the dissemination of the Alpha variant introduced heterogeneous detection into the equations. The agent-based model, initially developed to study influenza has rapidly been updated to study digital contact tracing [59], and then reactive vaccination against COVID-19. These works have benefited from the unprecedented amount of data produced during the pandemic [178]. Beside the implementation of monitoring and reporting systems, the use of this data has been made possible by a massive effort to gather, share and make this data easily accessible. At the onset of the pandemic, researchers had to manually aggregate data from available reports [17, 18] to analyse data on the first cases. In addition, the emergence of new variants has prompted the development of protocols for extensive genomic surveillance. Many authorities implemented data repositories and dashboards that allowed for quick access to the number of cases recorded per day. In most countries, governmental agencies provided platforms to share data about daily detected cases, hospitalisations and deaths, as in New York [179], France [42], Germany [180] or the

UK [181]. Besides national sharing, academic and public health organisations have aggregated this data to make it easier to access and enable comparison between countries. Among others, we can cite the COVID-19 Data Repository by the Center for Systems Science and Engineering by Johns Hopkins University [182], the OurWorldInData project [73], the Global;health project [183] or the platform provided by WHO [184]. For genomic surveillance, the GISAID database enabled the massive sharing of COVID-19 sequences and associated metadata. While it has been active since 2008, the GISAID initiative played a massive role in research during the COVID-19 pandemic. The high quality and massive amount of genomic data played a crucial role in the development of diagnostic tests and the first vaccines and enabled the monitoring in realtime of VOC [185]. In addition to epidemiological data, complex models need to be informed by data on the various factors contributing to the spread of the epidemic. Publicly available results of surveys on statistics of contact tracing (number of contacts traced, delay) [57], social distancing [65] provided useful data to enriched models on contact tracing and reduction of presence in different settings for epidemic assessment and scenario analysis. Contact rates and their reduction due to NPIs were not directly accessible. The average number of contacts per individual was quantified before the pandemic through surveys [116]. To access variations in contact rates, we used mobility data recorded by Google [165], which provided a relative change in the number of people travelling to a specific setting compared to before the pandemic. We used this data as a proxy of the reduction of contact by setting. Other proxies have been widely used to estimate the impact of NPIs, mostly from digital data: mobile phones, google and social networks data, and GPS [186]. To prepare for future pandemics, it is necessary to continue developing tools that allow for data sharing and accessibility. Even though the COVID-19 pandemic is behind us, it is critical to continue to develop, finance, and test agile and effective data sharing platforms to respond to future challenges. Concurrently, efforts should be made to continue to develop tools to collecting data on factors affecting the course of an epidemic: real-time contact matrix, networks of contact, socio-demographic data and traveller flows. Working on these issues in peacetime has two objectives. Firstly, it allows us to work outside the emergency of the pandemic, to carry out in-depth retrospective studies on the available data. Secondly, having data outside the pandemic period, for example on contacts or mobility, provides a baseline to compute how these factors change during a pandemic.

Bibliography

1. Li Qun, Guan Xuhua, Wu Peng, *et al.* Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus–Infected Pneumonia. *New England Journal of Medicine* **382,** 1199–1207. https://www.nejm.org/doi/full/10.1056/NEJMoa2001316 (2020).
2. Chen, Z., Azman, A. S., Chen, X., *et al.* Global landscape of SARS-CoV-2 genomic surveillance and data sharing. *Nature Genetics* **54,** 499–507. https://www.nature.com/articles/s41588-022-01033-y (2022).
3. Tian, H., Liu, Y., Li, Y., *et al.* An investigation of transmission control measures during the first 50 days of the COVID-19 epidemic in China. *Science* **368,** 638–642. https://www.science.org/doi/10.1126/science.abb6105 (2020).
4. *More EU nations ban travel from UK, fearing virus variant* 2020. https://apnews.com/article/europe-england-united-kingdom-coronavirus-pandemic-germany-06a35eb02ce885ae7b80b2e1851ab893.
5. Faucher, B., Sabbatini, C. E., Czuppon, P., *et al.* Drivers and impact of the early silent invasion of SARS-CoV-2 Alpha. *Nature Communications* **15,** 2152. https://www.nature.com/articles/s41467-024-46345-1 (2024).
6. Faucher, B., Assab, R., Roux, J., *et al.* Agent-based modelling of reactive vaccination of workplaces and schools against COVID-19. *Nature Communications* **13,** 1414. https://www.nature.com/articles/s41467-022-29015-y (2022).
7. *Covid-19 : face aux variants émergents, vacciner de manière réactive* https://www.has-sante.fr/jcms/p_3273017/fr/covid-19-face-aux-variants-emergents-vacciner-de-maniere-reactive.
8. Czuppon, P., Schertzer, E., Blanquart, F. & Débarre, F. The stochastic dynamics of early epidemics: probability of establishment, initial growth rate, and infection cluster size at first detection. *Journal of The Royal Society Interface* **18,** 20210575. https://royalsocietypublishing.org/doi/10.1098/rsif.2021.0575 (2021).
9. Pullano, G., Di Domenico, L., Sabbatini, C. E., *et al.* Underdetection of cases of COVID-19 in France threatens epidemic control. *Nature* **590,** 134–139. https://www.nature.com/articles/s41586-020-03095-6 (2021).
10. Domenico, L. D., Sabbatini, C. E., Pullano, G., Lévy-Bruhl, D. & Colizza, V. Impact of January 2021 curfew measures on SARS-CoV-2 B.1.1.7 circulation in France. *Eurosurveillance* **26,** 2100272. https://www.eurosurveillance.org/content/10.2807/1560-7917.ES.2021.26.15.2100272 (2021).
11. Di Domenico, L., Sabbatini, C. E., Boëlle, P.-Y., *et al.* Adherence and sustainability of interventions informing optimal control against the COVID-19 pandemic. *Communications Medicine* **1,** 1–13. https://www.nature.com/articles/s43856-021-00057-5 (2021).
12. Liu, Y., Ning, Z., Chen, Y., *et al.* Aerodynamic analysis of SARS-CoV-2 in two Wuhan hospitals. *Nature* **582,** 557–560. https://www.nature.com/articles/s41586-020-2271-3 (2020).

13. *Report of the WHO-China Joint Mission on Coronavirus Disease 2019 (COVID-19)* https://www.who.int/publications-detail-redirect/report-of-the-who-china-joint-mission-on-coronavirus-disease-2019-(covid-19).
14. Salje, H., Tran Kiem, C., Lefrancq, N., *et al.* Estimating the burden of SARS-CoV-2 in France. *Science* **369,** 208–211. https://www.science.org/doi/10.1126/science.abc3517 (2020).
15. Wu, F., Zhao, S., Yu, B., *et al.* A new coronavirus associated with human respiratory disease in China. *Nature* **579,** 265–269. https://www.nature.com/articles/s41586-020-2008-3 (2020).
16. *COVID-19 – Chronologie de l'action de l'OMS* https://www.who.int/fr/news/item/27-04-2020-who-timeline---covid-19.
17. Chinazzi, M., Davis, J. T., Ajelli, M., *et al.* The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. *Science* **368,** 395–400. https://www.science.org/doi/10.1126/science.aba9757 (2020).
18. Pinotti, F., Domenico, L. D., Ortega, E., *et al.* Tracing and analysis of 288 early SARS-CoV-2 infections outside China: A modeling study. *PLOS Medicine* **17,** e1003193. https://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.1003193 (2020).
19. *Coronavirus : circulation des variants du SARS-CoV-2* https://www.santepubliquefrance.fr/dossiers/coronavirus-covid-19/coronavirus-circulation-des-variants-du-sars-cov-2.
20. Xie, Z., Qin, Y., Li, Y., *et al.* Spatial and temporal differentiation of COVID-19 epidemic spread in mainland China and its influencing factors. *Science of The Total Environment* **744,** 140929. https://www.sciencedirect.com/science/article/pii/S0048969720344582 (2020).
21. Li, J., Lai, S., Gao, G. F. & Shi, W. The emergence, genomic diversity and global spread of SARS-CoV-2. *Nature* **600,** 408–418. https://www.nature.com/articles/s41586-021-04188-6 (2021).
22. Yang, J., Li, J., Lai, S., *et al.* Uncovering two phases of early intercontinental COVID-19 transmission dynamics. *Journal of Travel Medicine* **27,** taaa200. https://doi.org/10.1093/jtm/taaa200 (2020).
23. Worobey, M., Pekar, J., Larsen, B. B., *et al.* The emergence of SARS-CoV-2 in Europe and North America. *Science* **370,** 564–570. https://www.science.org/doi/10.1126/science.abc8169 (2020).
24. Lemey, P., Hong, S. L., Hill, V., *et al.* Accommodating individual travel history and unsampled diversity in Bayesian phylogeographic inference of SARS-CoV-2. *Nature Communications* **11,** 5110. https://www.nature.com/articles/s41467-020-18877-9 (2020).
25. Pullano, G., Pinotti, F., Valdano, E., *et al.* Novel coronavirus (2019-nCoV) early-stage importation risk to Europe, January 2020. *Eurosurveillance* **25,** 2000057. https://www.eurosurveillance.org/content/10.2807/1560-7917.ES.2020.25.4.2000057 (2020).
26. Spiteri, G., Fielding, J., Diercke, M., *et al.* First cases of coronavirus disease 2019 (COVID-19) in the WHO European Region, 24 January to 21 February 2020. *Eurosurveillance* **25,** 2000178. https://www.eurosurveillance.org/content/10.2807/1560-7917.ES.2020.25.9.2000178 (2020).

27. *Genetic Variants of SARS-CoV-2—What Do They Mean? | Vaccination | JAMA | JAMA Network* https://jamanetwork.com/journals/jama/fullarticle/2775006.

28. Rambaut, A., Holmes, E. C., O'Toole, *et al.* A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. *Nature Microbiology* **5,** 1403–1407. https://www.nature.com/articles/s41564-020-0770-5 (2020).

29. CDC. *Coronavirus Disease 2019 (COVID-19)* 2020. https://www.cdc.gov/coronavirus/2019-ncov/variants/variant-classifications.html.

30. *SARS-CoV-2 variants of concern as of 5 January 2024* 2021. https://www.ecdc.europa.eu/en/covid-19/variants-concern.

31. *Tracking SARS-CoV-2 variants* https://www.who.int/activities/tracking-SARS-CoV-2-variants.

32. *WHO announces simple, easy-to-say labels for SARS-CoV-2 Variants of Interest and Concern* https://www.who.int/news/item/31-05-2021-who-announces-simple-easy-to-say-labels-for-sars-cov-2-variants-of-interest-and-concern.

33. Volz, E., Hill, V., McCrone, J. T., *et al.* Evaluating the Effects of SARS-CoV-2 Spike Mutation D614G on Transmissibility and Pathogenicity. *Cell* **184,** 64–75.e11. https://www.sciencedirect.com/science/article/pii/S0092867420315373 (2021).

34. O'Toole, Hill, V., Pybus, O. G., *et al.* Tracking the international spread of SARS-CoV-2 lineages B.1.1.7 and B.1.351/501Y-V2 with grinch. *Wellcome Open Research* **6,** 121. https://doi.org/10.12688/wellcomeopenres.16661.2 (2021).

35. *Coronavirus : chiffres clés et évolution de la COVID-19 en France et dans le Monde* https://www.santepubliquefrance.fr/dossiers/coronavirus-covid-19/coronavirus-chiffres-cles-et-evolution-de-la-covid-19-en-france-et-dans-le-monde.

36. *Estimated transmissibility and impact of SARS-CoV-2 lineage B.1.1.7 in England | Science* https://www.science.org/doi/10.1126/science.abg3055.

37. Tegally, H., Wilkinson, E., Giovanetti, M., *et al.* Detection of a SARS-CoV-2 variant of concern in South Africa. *Nature* **592,** 438–443. https://www.nature.com/articles/s41586-021-03402-9 (2021).

38. Tegally, H., Wilkinson, E., Tsui, J. L. H., *et al.* Dispersal patterns and influence of air travel during the global expansion of SARS-CoV-2 variants of concern. *Cell* **186,** 3277–3290.e16. https://www.sciencedirect.com/science/article/pii/S0092867423006414 (2023).

39. Kraemer, M. U. G., Hill, V., Ruis, C., *et al.* Spatiotemporal invasion dynamics of SARS-CoV-2 lineage B.1.1.7 emergence. *Science* **373,** 889–895. https://www.science.org/doi/10.1126/science.abj0113 (2021).

40. Dellicour, S., Durkin, K., Hong, S. L., *et al.* A Phylodynamic Workflow to Rapidly Gain Insights into the Dispersal History and Dynamics of SARS-CoV-2 Lineages. *Molecular Biology and Evolution* **38,** 1608–1613. https://doi.org/10.1093/molbev/msaa284 (2021).

41. Gräf, T., Martinez, A. A., Bello, G., *et al.* Dispersion patterns of SARS-CoV-2 variants Gamma, Lambda and Mu in Latin America and the Caribbean. *Nature Communications* **15,** 1837. https://www.nature.com/articles/s41467-024-46143-9 (2024).

42. *Tableau de bord COVID-19 Suivi de l'épidémie de COVID-19 en France* https://dashboard.covid19.data.gouv.fr/vue-d-ensemble?location=FRA.

43. Shu, Y. & McCauley, J. GISAID: Global initiative on sharing all influenza data – from vision to reality. *Eurosurveillance* **22,** 30494. https://www.eurosurveillance.org/content/10.2807/1560-7917.ES.2017.22.13.30494 (2017).

44. *Strategies for the surveillance of COVID-19* 2020. https://www.ecdc.europa.eu/en/publications-data/strategies-surveillance-covid-19.

45. Ibrahim, N. K. Epidemiologic surveillance for controlling Covid-19 pandemic: types, challenges and implications. *Journal of Infection and Public Health* **13,** 1630–1638. https://www.sciencedirect.com/science/article/pii/S1876034120306031 (2020).

46. Han, A. X., Toporowski, A., Sacks, J. A., *et al.* SARS-CoV-2 diagnostic testing rates determine the sensitivity of genomic surveillance programs. *Nature Genetics* **55,** 26–33. https://www.nature.com/articles/s41588-022-01267-w (2023).

47. *Enquêtes Flash : évaluation de la circulation des variants du SARS-CoV-2 en France* https://www.santepubliquefrance.fr/etudes-et-enquetes/enquetes-flash-evaluation-de-la-circulation-des-variants-du-sars-cov-2-en-france.

48. *Consortium EMERGEN* https://www.santepubliquefrance.fr/dossiers/coronavirus-covid-19/consortium-emergen.

49. Mendelson, M., Venter, F., Moshabela, M., *et al.* The political theatre of the UK's travel ban on South Africa. *The Lancet* **398,** 2211–2213. https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(21)02752-5/fulltext (2021).

50. *Could the Delta variant cancel European travel this summer?* 2021. https://www.euronews.com/travel/2021/06/17/will-the-delta-variant-affect-travel-restrictions-in-europe.

51. Kucharski, A. J., Jit, M., Logan, J. G., *et al.* Travel measures in the SARS-CoV-2 variant era need clear objectives. *The Lancet* **399,** 1367–1369. https://www.sciencedirect.com/science/article/pii/S014067362200366X (2022).

52. Bajardi, P., Poletto, C., Ramasco, J. J., *et al.* Human Mobility Networks, Travel Restrictions, and the Global Spread of 2009 H1N1 Pandemic. *PLOS ONE* **6,** e16591. https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0016591 (2011).

53. Poletto, C., Gomes, M. F., Piontti, A. P. y., *et al.* Assessing the impact of travel restrictions on international spread of the 2014 West African Ebola epidemic. *Eurosurveillance* **19,** 20936. https://www.eurosurveillance.org/content/10.2807/1560-7917.ES2014.19.42.20936 (2014).

54. Grépin, K. A., Ho, T.-L., Liu, Z., *et al.* Evidence of the effectiveness of travel-related measures during the early phase of the COVID-19 pandemic: a rapid systematic review. *BMJ Global Health* **6,** e004537. https://gh.bmj.com/content/6/3/e004537 (2021).

55. Russell, T. W., Wu, J. T., Clifford, S., *et al.* Effect of internationally imported cases on internal spread of COVID-19: a mathematical modelling study. *The Lancet Public Health* **6,** e12–e20. https://www.thelancet.com/journals/lanpub/article/PIIS2468-2667(20)30263-2/fulltext (2021).

56. Pozo-Martin, F., Beltran Sanchez, M. A., Müller, S. A., *et al.* Comparative effectiveness of contact tracing interventions in the context of the COVID-19 pandemic: a systematic review. *European Journal of Epidemiology* **38,** 243–266. https://doi.org/10.1007/s10654-023-00963-z (2023).

57. *Contact-tracing* https://www.santepubliquefrance.fr/dossiers/coronavirus-covid-19/contact-tracing.

58. Quilty, B. J., Clifford, S., Hellewell, J., *et al.* Quarantine and testing strategies in contact tracing for SARS-CoV-2: a modelling study. *The Lancet Public Health* **6,** e175–e183. https://www.thelancet.com/journals/lanpub/article/PIIS2468-2667(20)30308-X/fulltext (2021).

59. López, J. A. M., García, B. A., Bentkowski, P., *et al.* Anatomy of digital contact tracing: Role of age, transmission setting, adoption, and case detection. *Science Advances.* https://www.science.org/doi/abs/10.1126/sciadv.abd8750 (2021).

60. Hale, T., Angrist, N., Goldszmidt, R., *et al.* A global panel database of pandemic policies (Oxford COVID-19 Government Response Tracker). *Nature Human Behaviour* **5,** 529–538. https://www.nature.com/articles/s41562-021-01079-8 (2021).

61. *Décret n° 2020-260 du 16 mars 2020 portant réglementation des déplacements dans le cadre de la lutte contre la propagation du virus covid-19* 2020.

62. *Décret n° 2020-1310 du 29 octobre 2020 prescrivant les mesures générales nécessaires pour faire face à l'épidémie de covid-19 dans le cadre de l'état d'urgence sanitaire* 2020.

63. *Décret n° 2020-545 du 11 mai 2020 prescrivant les mesures générales nécessaires pour faire face à l'épidémie de covid-19 dans le cadre de l'état d'urgence sanitaire* 2020.

64. *Décret n° 2021-1725 du 21 décembre 2021 modifiant les conditions de mise en œuvre du télétravail dans la fonction publique et la magistrature - Légifrance* https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000044538035.

65. *CoviPrev : une enquête pour suivre l'évolution des comportements et de la santé mentale pendant l'épidémie de COVID-19* https://www.santepubliquefrance.fr/etudes-et-enquetes/coviprev-une-enquete-pour-suivre-l-evolution-des-comportements-et-de-la-sante-mentale-pendant-l-epidemie-de-covid-19.

66. Lison, A., Banholzer, N., Sharma, M., *et al.* Effectiveness assessment of non-pharmaceutical interventions: lessons learned from the COVID-19 pandemic. *The Lancet Public Health* **8,** e311–e317. https://www.thelancet.com/journals/lanpub/article/PIIS2468-2667(23)00046-4/fulltext (2023).

67. Flaxman, S., Mishra, S., Gandy, A., *et al.* Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. *Nature* **584,** 257–261. https://www.nature.com/articles/s41586-020-2405-7 (2020).

68. Brauner, J. M., Mindermann, S., Sharma, M., *et al.* Inferring the effectiveness of government interventions against COVID-19. *Science* **371,** eabd9338. https://www.science.org/doi/10.1126/science.abd9338 (2021).

69. Banholzer, N., Weenen, E. v., Lison, A., *et al.* Estimating the effects of non-pharmaceutical interventions on the number of new infections with COVID-19 during the first epidemic wave. *PLOS ONE* **16,** e0252827. https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0252827 (2021).

70. Polack, F. P., Thomas, S. J., Kitchin, N., *et al.* Safety and Efficacy of the BNT162b2 mRNA Covid-19 Vaccine. *New England Journal of Medicine* **383,** 2603–2615. https://doi.org/10.1056/NEJMoa2034577 (2020).

71. Wu, N., Joyal-Desmarais, K., Ribeiro, P. A. B., *et al.* Long-term effectiveness of COVID-19 vaccines against infections, hospitalisations, and mortality in adults: findings from a rapid living systematic evidence synthesis and meta-analysis up to December, 2022. *The Lancet Respiratory Medicine* **11,** 439–452. https://www.thelancet.com/journals/lanres/article/PIIS2213-2600(23)00015-2/fulltext (2023).

72. Link-Gelles, R. Early Estimates of Updated 2023–2024 (Monovalent XBB.1.5) COVID-19 Vaccine Effectiveness Against Symptomatic SARS-CoV-2 Infection Attributable to Co-Circulating Omicron Variants Among Immunocompetent Adults — Increasing Community Access to Testing Program, United States, September 2023–January 2024. *MMWR. Morbidity and Mortality Weekly Report* **73.** https://www.cdc.gov/mmwr/volumes/73/wr/mm7304a2.htm (2024).

73. Mathieu, E., Ritchie, H., Rodés-Guirao, L., *et al.* Coronavirus Pandemic (COVID-19). *Our World in Data.* https://ourworldindata.org/covid-vaccinations (2020).

74. *Logistique vaccinale* https://www.santepubliquefrance.fr/dossiers/coronavirus-covid-19/logistique-vaccinale.

75. *Vaccination contre la COVID-19 : un défi logistique et scientifique* https://www.santepubliquefrance.fr/presse/2020/vaccination-contre-la-covid-19-un-defi-logistique-et-scientifique.

76. Kiem, C. T., Massonnaud, C. R., Levy-Bruhl, D., *et al.* A modelling study investigating short and medium-term challenges for COVID-19 vaccination: From prioritisation to the relaxation of measures. *eClinicalMedicine* **38.** https://www.thelancet.com/journals/eclinm/article/PIIS2589-5370(21)00281-9/fulltext (2021).

77. Rousseau, R., Garcia-Zorita, C. & Sanz-Casado, E. Publications during COVID-19 times: An unexpected overall increase. *Journal of Informetrics* **17,** 101461. https://www.sciencedirect.com/science/article/pii/S175115772300086X (2023).

78. Becker, A. D., Grantz, K. H., Hegde, S. T., *et al.* Development and dissemination of infectious disease dynamic transmission models during the COVID-19 pandemic: what can we learn from other pathogens and how can we move forward? *The Lancet Digital Health* **3,** e41–e50. https://www.thelancet.com/journals/landig/article/PIIS2589-7500(20)30268-5/fulltext (2021).

79. Khan, K., Sears, J., Hu, V. W., *et al.* Potential for the international spread of middle East respiratory syndrome in association with mass gatherings in saudi arabia. *PLoS currents* **5,** ecurrents.outbreaks.a7b70897ac2fa4f79b59f90d24c860b8. https://doi.org/10.1371/currents.outbreaks.a7b70897ac2fa4f79b59f90d24c860b8 (2013).

80. Poletto, C., Boëlle, P.-Y. & Colizza, V. Risk of MERS importation and onward transmission: a systematic review and analysis of cases reported to WHO. *BMC Infectious Diseases* **16,** 448. https://doi.org/10.1186/s12879-016-1787-5 (2016).

81. Bogoch, I. I., Watts, A., Thomas-Bachli, A., *et al.* Potential for global spread of a novel coronavirus from China. *Journal of Travel Medicine* **27,** taaa011. https://doi.org/10.1093/jtm/taaa011 (2020).

82. Gilbert, M., Pullano, G., Pinotti, F., *et al.* Preparedness and vulnerability of African countries against importations of COVID-19: a modelling study. *The Lancet* **395,** 871–877. https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(20)30411-6/fulltext (2020).

83. Lai, S., Floyd, J. & Tatem, A. Preliminary risk analysis of the international spread of new COVID-19 variants, lineage B.1.1.7, B.1.351 and P.

84. Du, Z., Wang, L., Yang, B., *et al.* Risk for International Importations of Variant SARS-CoV-2 Originating in the United Kingdom - Volume 27, Number 5—May 2021 - Emerging Infectious Diseases journal - CDC. https://wwwnc.cdc.gov/eid/article/27/5/21-0050_article.

85. *Report 1 - Estimating the potential total number of novel Coronavirus (2019-nCoV) cases in Wuhan City, China* https://www.imperial.ac.uk/medicine/departments/school-public-health/infectious-disease-epidemiology/mrc-global-infectious-disease-analysis/disease-areas/covid-19/report-1-case-estimates-of-covid-19/.

86. Niehus, R., Salazar, P. M. D., Taylor, A. R. & Lipsitch, M. Using observational data to quantify bias of traveller-derived COVID-19 prevalence estimates in Wuhan, China. *The Lancet Infectious Diseases* **20,** 803–808. https://www.thelancet.com/journals/laninf/article/PIIS1473-3099(20)30229-2/fulltext (2020).

87. Jijón, S., Czuppon, P., Blanquart, F. & Débarre, F. Using early detection data to estimate the date of emergence of an epidemic outbreak. *PLOS Computational Biology* **20,** e1011934. https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1011934 (2024).

88. Keeling, M. J. & Rohani, P. Estimating spatial coupling in epidemiological systems: a mechanistic approach. *Ecology Letters* **5,** 20–29. https://onlinelibrary.wiley.com/doi/abs/10.1046/j.1461-0248.2002.00268.x (2002).

89. May, R. M. & Anderson, R. M. Spatial heterogeneity and the design of immunization programs. *Mathematical Biosciences* **72,** 83–111. https://www.sciencedirect.com/science/article/pii/0025556484900634 (1984).

90. Colizza, V., Barrat, A., Barthélemy, M. & Vespignani, A. The role of the airline transportation network in the prediction and predictability of global epidemics. *Proceedings of the National Academy of Sciences* **103,** 2015–2020. https://www.pnas.org/doi/10.1073/pnas.0510525103 (2006).

91. Longini, I. M. A mathematical model for predicting the geographic spread of new infectious agents. *Mathematical Biosciences* **90,** 367–383. https://www.sciencedirect.com/science/article/pii/0025556488900752 (1988).

92. Grais, R. F., Hugh Ellis, J. & Glass, G. E. Assessing the impact of airline travel on the geographic spread of pandemic influenza. *European Journal of Epidemiology* **18,** 1065–1072. https://doi.org/10.1023/A:1026140019146 (2003).

93. Hufnagel, L., Brockmann, D. & Geisel, T. Forecast and control of epidemics in a globalized world. *Proceedings of the National Academy of Sciences* **101,** 15124–15129. https://www.pnas.org/doi/10.1073/pnas.0308344101 (2004).

94. Cooper, B. S., Pitman, R. J., Edmunds, W. J. & Gay, N. J. Delaying the International Spread of Pandemic Influenza. *PLOS Medicine* **3,** e212. https://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.0030212 (2006).

95. Hollingsworth, T. D., Ferguson, N. M. & Anderson, R. M. Will travel restrictions control the international spread of pandemic influenza? *Nature Medicine* **12,** 497–499. https://www.nature.com/articles/nm0506-497 (2006).

96. Germann, T. C., Kadau, K., Longini, I. M. & Macken, C. A. Mitigation strategies for pandemic influenza in the United States. *Proceedings of the National Academy of Sciences* **103,** 5935–5940. https://www.pnas.org/doi/full/10.1073/pnas.0601266103 (2006).

97. Colizza, V., Barrat, A., Barthelemy, M., Valleron, A.-J. & Vespignani, A. Modeling the Worldwide Spread of Pandemic Influenza: Baseline Case and Containment Interventions. *PLOS Medicine* **4,** e13. https://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.0040013 (2007).

98. *Vision | GLEAM Project* https://www.gleamproject.org/vision.

99. Balcan, D., Gonçalves, B., Hu, H., *et al.* Modeling the spatial spread of infectious diseases: The GLobal Epidemic and Mobility computational model. *Journal of Computational Science* **1,** 132–145. https://www.sciencedirect.com/science/article/pii/S1877750310000438 (2010).

100. Balcan, D., Colizza, V., Gonçalves, B., *et al.* Multiscale mobility networks and the spatial spreading of infectious diseases. *Proceedings of the National Academy of Sciences* **106,** 21484–21489. https://www.pnas.org/doi/abs/10.1073/pnas.0906910106 (2009).

101. Zhang, Q., Sun, K., Chinazzi, M., *et al.* Spread of Zika virus in the Americas. *Proceedings of the National Academy of Sciences* **114,** E4334–E4343. https://www.pnas.org/doi/abs/10.1073/pnas.1620161114 (2017).

102. Parino, F., Gustani-Buss, E., Bedford, T., *et al. Integrating dynamical modeling and phylogeographic inference to characterize global influenza circulation* 2024. https://www.medrxiv.org/content/10.1101/2024.03.14.24303719v1.

103. Davis, J. T., Chinazzi, M., Perra, N., *et al.* Cryptic transmission of SARS-CoV-2 and the first COVID-19 wave. *Nature* **600,** 127–132. https://www.nature.com/articles/s41586-021-04130-w (2021).

104. Gautreau, A., Barrat, A. & Barthélemy, M. Global disease spread: Statistics and estimation of arrival times. *Journal of Theoretical Biology* **251,** 509–522. https://www.sciencedirect.com/science/article/pii/S002251930700608X (2008).

105. Scalia Tomba, G. & Wallinga, J. A simple explanation for the low impact of border control as a countermeasure to the spread of an infectious disease. *Mathematical Biosciences. BICOMP 2008* **214,** 70–72. https://www.sciencedirect.com/science/article/pii/S0025556408000278 (2008).

106. Brockmann, D. & Helbing, D. The Hidden Geometry of Complex, Network-Driven Contagion Phenomena. *Science* **342,** 1337–1342. https://www.science.org/doi/10.1126/science.1245200 (2013).

107. Kermack, W. O., McKendrick, A. G. & Walker, G. T. A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* **115,** 700–721. https://royalsocietypublishing.org/doi/10.1098/rspa.1927.0118 (1997).

108. Keeling, M. & Ross, J. On methods for studying stochastic disease dynamics. *Journal of The Royal Society Interface* **5,** 171–181. https://royalsocietypublishing.org/doi/10.1098/rsif.2007.1106 (2007).

109. Schenzle, D. An age-structured model of pre- and post-vaccination measles transmission. *IMA journal of mathematics applied in medicine and biology* **1,** 169–191. https://doi.org/10.1093/imammb/1.2.169 (1984).

110. Wallinga, J., Teunis, P. & Kretzschmar, M. Using data on social contacts to estimate age-specific transmission parameters for respiratory-spread infectious agents. *American Journal of Epidemiology* **164,** 936–944. https://doi.org/10.1093/aje/kwj317 (2006).

111. Zhang, J., Litvinova, M., Liang, Y., *et al.* Changes in contact patterns shape the dynamics of the COVID-19 outbreak in China. *Science* **368,** 1481–1486. https://www.science.org/doi/10.1126/science.abb8001 (2020).

112. Lai, S., Ruktanonchai, N. W., Zhou, L., *et al.* Effect of non-pharmaceutical interventions to contain COVID-19 in China. *Nature* **585,** 410–413. https://www.nature.com/articles/s41586-020-2293-x (2020).

113. Kucharski, A. J., Russell, T. W., Diamond, C., *et al.* Early dynamics of transmission and control of COVID-19: a mathematical modelling study. *The Lancet Infectious Diseases* **20,** 553–558. https://www.thelancet.com/article/S1473-3099(20)30144-4/fulltext (2020).

114. Di Domenico, L., Pullano, G., Sabbatini, C. E., Boëlle, P.-Y. & Colizza, V. Impact of lockdown on COVID-19 epidemic in Île-de-France and possible exit strategies. *BMC Medicine* **18,** 240. https://doi.org/10.1186/s12916-020-01698-4 (2020).

115. Davies, N. G., Kucharski, A. J., Eggo, R. M., *et al.* Effects of non-pharmaceutical interventions on COVID-19 cases, deaths, and demand for hospital services in the UK: a modelling study. *The Lancet Public Health* **5,** e375–e385. https://www.thelancet.com/journals/lanpub/article/PIIS2468-2667(20)30133-X/fulltext (2020).

116. Béraud, G., Kazmercziak, S., Beutels, P., *et al.* The French Connection: The First Large Population-Based Contact Survey in France Relevant for the Spread of Infectious Diseases. *PLOS ONE* **10,** e0133203. https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0133203 (2015).

117. Sabbatini, C. E., Pullano, G., Di Domenico, L., *et al.* The impact of spatial connectivity on NPIs effectiveness. *BMC Infectious Diseases* **24,** 21. https://doi.org/10.1186/s12879-023-08900-x (2024).

118. Kiem, C. T., Crépey, P., Bosetti, P., *et al.* Lockdown as a last resort option in case of COVID-19 epidemic rebound: a modelling study. *Eurosurveillance* **26,** 2001536. https://www.eurosurveillance.org/content/10.2807/1560-7917.ES.2021.26.22.2001536 (2021).

119. Roux, J., Massonnaud, C. R., Colizza, V., Cauchemez, S. & Crépey, P. Modeling the impact of national and regional lockdowns on the 2020 spring wave of COVID-19 in France. *Scientific Reports* **13,** 1834. https://www.nature.com/articles/s41598-023-28687-w (2023).

120. Tran Kiem, C., Bosetti, P., Paireau, J., *et al.* SARS-CoV-2 transmission across age groups in France and implications for control. *Nature Communications* **12,** 6895. https://www.nature.com/articles/s41467-021-27163-1 (2021).

121. Bosetti, P., Kiem, C. T., Yazdanpanah, Y., *et al.* Impact of mass testing during an epidemic rebound of SARS-CoV-2: a modelling study using the example of France. *Eurosurveillance* **26,** 2001978. https://www.eurosurveillance.org/content/10.2807/1560-7917.ES.2020.26.1.2001978 (2021).

122. Mossong, J., Hens, N., Jit, M., *et al.* Social Contacts and Mixing Patterns Relevant to the Spread of Infectious Diseases. *PLOS Medicine* **5,** e74. https://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.0050074 (2008).

123. Contreras, D. A., Colosi, E., Bassignana, G., Colizza, V. & Barrat, A. Impact of contact data resolution on the evaluation of interventions in mathematical models of infectious diseases. *Journal of The Royal Society Interface* **19,** 20220164. https://royalsocietypublishing.org/doi/10.1098/rsif.2022.0164 (2022).

124. Durrett, R. Stochastic Spatial Models. *SIAM Review* **41,** 677–718. https://epubs.siam.org/doi/abs/10.1137/S0036144599354707 (1999).

125. Britton, T. & O'neill, P. D. Bayesian Inference for Stochastic Epidemics in Populations with Random Social Structure. *Scandinavian Journal of Statistics* **29,** 375–390. https://onlinelibrary.wiley.com/doi/abs/10.1111/1467-9469.00296 (2002).

126. Meyers, L. A., Pourbohloul, B., Newman, M. E. J., Skowronski, D. M. & Brunham, R. C. Network theory and SARS: predicting outbreak diversity. *Journal of Theoretical Biology* **232,** 71–81. https://www.sciencedirect.com/science/article/pii/S0022519304003510 (2005).

127. Colosi, E., Bassignana, G., Barrat, A., *et al.* Minimising school disruption under high incidence conditions due to the Omicron variant in France, Switzerland, Italy, in January 2022. *Eurosurveillance* **28,** 2200192. https://www.eurosurveillance.org/content/10.2807/1560-7917.ES.2023.28.5.2200192 (2023).

128. Colosi, E., Bassignana, G., Contreras, D. A., *et al.* Screening and vaccination against COVID-19 to minimise school closure: a modelling study. *The Lancet Infectious Diseases* **22,** 977–989. https://www.thelancet.com/journals/laninf/article/PIIS1473-3099(22)00138-4/fulltext (2022).

129. Willem, L., Verelst, F., Bilcke, J., Hens, N. & Beutels, P. Lessons from a decade of individual-based models for infectious disease transmission: a systematic review (2006-2015). *BMC Infectious Diseases* **17,** 612. https://doi.org/10.1186/s12879-017-2699-8 (2017).

130. Ferguson, N. M., Cummings, D. A. T., Cauchemez, S., *et al.* Strategies for containing an emerging influenza pandemic in Southeast Asia. *Nature* **437,** 209–214. https://www.nature.com/articles/nature04017 (2005).

131. Longini, I. M., Nizam, A., Xu, S., *et al.* Containing Pandemic Influenza at the Source. *Science* **309,** 1083–1087. https://www.science.org/doi/10.1126/science.1115717 (2005).

132. Eubank, S., Guclu, H., Anil Kumar, V. S., *et al.* Modelling disease outbreaks in realistic urban social networks. *Nature* **429,** 180–184. https://www.nature.com/articles/nature02541 (2004).

133. Maire, N., Shillcutt, S. D., Walker, D. G., Tediosi, F. & Smith, T. A. Cost-Effectiveness of the Introduction of a Pre-Erythrocytic Malaria Vaccine into the Expanded Program on Immunization in Sub-Saharan Africa: Analysis of Uncertainties Using a Stochastic Individual-Based Simulation Model of *Plasmodium falciparum* Malaria. *Value in Health* **14,** 1028–1038. https://www.sciencedirect.com/science/article/pii/S1098301511015269 (2011).

134. Poletti, P., Merler, S., Ajelli, M., *et al.* Evaluating vaccination strategies for reducing infant respiratory syncytial virus infection in low-income settings. *BMC Medicine* **13,** 49. https://doi.org/10.1186/s12916-015-0283-x (2015).

135. Aleta, A., Martín-Corral, D., Pastore y Piontti, A., *et al.* Modelling the impact of testing, contact tracing and household quarantine on second waves of COVID-19. *Nature Human Behaviour* **4,** 964–971. https://www.nature.com/articles/s41562-020-0931-9 (2020).

136. Kucharski, A. J., Klepac, P., Conlan, A. J. K., *et al.* Effectiveness of isolation, testing, contact tracing, and physical distancing on reducing transmission of SARS-CoV-2 in different settings: a mathematical modelling study. *The Lancet Infectious Diseases* **20,** 1151–1160. https://www.thelancet.com/article/S1473-3099(20)30457-6/fulltext (2020).

137. Sun, K., Wang, W., Gao, L., *et al.* Transmission heterogeneities, kinetics, and controllability of SARS-CoV-2. *Science* **371,** eabe2424. https://www.science.org/doi/10.1126/science.abe2424 (2021).

138. Crépey, P., Noël, H. & Alizon, S. Challenges for mathematical epidemiological modelling. *Anaesthesia, Critical Care & Pain Medicine* **41,** 101053. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8902375/ (2022).

139. Sofonea, M. T., Cauchemez, S. & Boëlle, P.-Y. Epidemic models: why and how to use them. *Anaesthesia, Critical Care & Pain Medicine* **41,** 101048. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8882476/ (2022).

140. Zelner, J., Riou, J., Etzioni, R. & Gelman, A. Accounting for uncertainty during a pandemic. *Patterns* **2.** https://www.cell.com/patterns/abstract/S2666-3899(21)00153-7 (2021).

141. Brito, A. F., Semenova, E., Dudas, G., *et al.* Global disparities in SARS-CoV-2 genomic surveillance. *Nature Communications* **13,** 7003. https://www.nature.com/articles/s41467-022-33713-y (2022).

142. *Investigation of SARS-CoV-2 variants of concern: technical briefings* 2021. https://www.gov.uk/government/publications/investigation-of-novel-sars-cov-2-variant-variant-of-concern-20201201.

143. *IATA - Home* https://www.iata.org/en/.

144. *Sequencing of SARS-CoV-2 - first update* 2021. https://www.ecdc.europa.eu/en/publications-data/sequencing-sars-cov-2.

145. *Detection and characterisation capability and capacity for SARS-CoV-2 variants within the EU/EEA* 2021. https://www.ecdc.europa.eu/en/publications-data/detection-and-characterisation-capability-and-capacity-sars-cov-2-variants.

146. *Genomic sequencing of SARS-CoV-2: a guide to implementation for maximum impact on public health* https://www.who.int/publications-detail-redirect/9789240018440.

147. Plotkin, S. A. Vaccines: past, present and future. *Nature Medicine* **11,** S5–S11. https://www.nature.com/articles/nm1209 (2005).

148. Orenstein, W. A. & Ahmed, R. Simply put: Vaccination saves lives. *Proceedings of the National Academy of Sciences* **114,** 4031–4033. https://www.pnas.org/doi/full/10.1073/pnas.1704507114 (2017).

149. *"Herd Immunity": A Rough Guide | Clinical Infectious Diseases | Oxford Academic* https://academic.oup.com/cid/article/52/7/911/299077.

150. *Covid-19: More variant hotspots to get surge tests and jabs* https://www.bbc.com/news/uk-57172139.

151. *Cologne aims to vaccinate urban hot spots – DW – 05/08/2021* https://www.dw.com/en/covid-cologne-project-aims-to-vaccinate-urban-hot-spots/a-57472989.

152. *COVID-19 : PRESENCE DU VARIANT INDIEN DANS L'EUROMETROPOLE ET PLAN D'ACTIONS IMMEDIAT* tech. rep. (). https://www.grand-est.ars.sante.fr/system/files/2021-06/CP%20Variant%20Indien.pdf.

153. Grais, R., Conlan, A., Ferrari, M., *et al.* Time is of the essence: exploring a measles outbreak response vaccination in Niamey, Niger. *Journal of The Royal Society Interface* **5,** 67–74. https://royalsocietypublishing.org/doi/full/10.1098/rsif.2007.1038 (2007).

154. Henao-Restrepo, A. M., Longini, I. M., Egger, M., *et al.* Efficacy and effectiveness of an rVSV-vectored vaccine expressing Ebola surface glycoprotein: interim results from the Guinea ring vaccination cluster-randomised trial. *The Lancet* **386,** 857–866. https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(15)61117-5/fulltext (2015).

155. Deen, J. & von Seidlein, L. The case for ring vaccinations with special consideration of oral cholera vaccines. *Human Vaccines & Immunotherapeutics* **14,** 2069–2074. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6149944/ (2018).

156. Geddes, A. M. The history of smallpox. *Clinics in Dermatology* **24,** 152–157. https://www.sciencedirect.com/science/article/pii/S0738081X05001707 (2006).

157. Merler, S., Ajelli, M., Fumanelli, L., *et al.* Containing Ebola at the Source with Ring Vaccination. *PLOS Neglected Tropical Diseases* **10,** e0005093. https://journals.plos.org/plosntds/article?id=10.1371/journal.pntd.0005093 (2016).

158. Pegorie, M., Shankar, K., Welfare, W. S., *et al.* Measles outbreak in Greater Manchester, England, October 2012 to September 2013: epidemiology and control. *Eurosurveillance* **19,** 20982. https://www.eurosurveillance.org/content/10.2807/1560-7917.ES2014.19.49.20982 (2014).

159. Ravindra, K., Malik, V. S., Padhi, B. K., Goel, S. & Gupta, M. Asymptomatic infection and transmission of COVID-19 among clusters: systematic review and meta-analysis. *Public Health* **203,** 100–109. https://www.sciencedirect.com/science/article/pii/S0033350621004649 (2022).

160. Casey-Bryars, M., Griffin, J., McAloon, C., *et al.* Presymptomatic transmission of SARS-CoV-2 infection: a secondary analysis using published data. *BMJ Open* **11,** e041240. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8245290/ (2021).

161. SPF. *COVID-19 : point épidémiologique du 6 mai 2021* https://www.santepubliquefrance.fr/maladies-et-traumatismes/maladies-et-infections-respiratoires/infection-a-coronavirus/documents/bulletin-national/covid-19-point-epidemiologique-du-6-mai-2021.

162. Bubar, K. M., Reinholt, K., Kissler, S. M., *et al.* Model-informed COVID-19 vaccine prioritization strategies by age and serostatus. *Science* **371,** 916–921. https://www.science.org/doi/10.1126/science.abe6959 (2021).

163. Matrajt, L., Eaton, J., Leung, T. & Brown, E. R. Vaccine optimization for COVID-19: Who to vaccinate first? *Science Advances* **7,** eabf1374. https://www.science.org/doi/10.1126/sciadv.abf1374 (2021).

164. Di Domenico, L. & Colizza, V. *Impact of the Timing of the Booster Dose Against Sars-Cov-2 Omicron Variant in France in Winter-Spring 2021-2022* SSRN Scholarly Paper. Rochester, NY, 2023. https://papers.ssrn.com/abstract=4652917.

165. *google mobility data - Recherche Google* https://www.google.com/search?q=google+mobility+data&oq=google+mobility+data&aqs=chrome..69i57j0i22i30l9.3715j0j4&sourceid=chrome&ie=UTF-8.

166. Kang, M., Xin, H., Yuan, J., *et al.* Transmission dynamics and epidemiological characteristics of SARS-CoV-2 Delta variant infections in Guangdong, China, May to June 2021. *Eurosurveillance* **27,** 2100815. https://www.eurosurveillance.org/content/10.2807/1560-7917.ES.2022.27.10.2100815 (2022).

167. Davies, N. G., Klepac, P., Liu, Y., *et al.* Age-dependent effects in the transmission and control of COVID-19 epidemics. *Nature Medicine* **26,** 1205–1211. https://www.nature.com/articles/s41591-020-0962-9 (2020).

168. Krymova, E., Béjar, B., Thanou, D., *et al.* Trend estimation and short-term forecasting of COVID-19 cases and deaths worldwide. *Proceedings of the National Academy of Sciences* **119,** e2112656119. https://www.pnas.org/doi/full/10.1073/pnas.2112656119 (2022).

169. *COVID-19: Willingness to be vaccinated | YouGov* https://yougov.co.uk/international/articles/33674-covid-19-willingness-be-vaccinated?redirect_from=%2Ftopics%2Finternational%2Farticles-reports%2F2021%2F01%2F12%2Fcovid-19-willingness-be-vaccinated.

170. *Covid-19 : un pari épidémiologique perdu ?* https://www.larevuedupraticien.fr/article/covid-19-un-pari-epidemiologique-perdu.

171. Ward, J. K., Peretti-Watel, P., Bocquier, A., Seror, V. & Verger, P. Vaccine hesitancy and coercion: all eyes on France. *Nature Immunology* **20,** 1257–1259. https://www.nature.com/articles/s41590-019-0488-9 (2019).

172. Lazarus, J. V., Wyka, K., White, T. M., *et al.* Revisiting COVID-19 vaccine hesitancy around the world using data from 23 countries in 2021. *Nature Communications* **13,** 3801. https://www.nature.com/articles/s41467-022-31441-x (2022).

173. Sacre, A., Bambra, C., Wildman, J. M., *et al.* Socioeconomic inequalities in vaccine uptake: A global umbrella review. *PLOS ONE* **18,** e0294688. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10718431/ (2023).

174. Débarre, F., Lecoeur, E., Guimier, L., Jauffret-Roustide, M. & Jannot, A.-S. The French Covid-19 vaccination policy did not solve vaccination inequities: a nationwide study on 64.5 million people. *European Journal of Public Health* **32,** 825–830. https://doi.org/10.1093/eurpub/ckac125 (2022).

175. *Bilan du dispositif d'Aller vers* https://www.assurance-maladie.ameli.fr/qui-sommes-nous/publications-reference/assurance-maladie/rapports-thematiques/dispositif-aller-vers.

176. Mandel, A. & Veetil, V. The Economic Cost of COVID Lockdowns: An Out-of-Equilibrium Analysis. *Economics of Disasters and Climate Change* **4,** 431–451. https://doi.org/10.1007/s41885-020-00066-z (2020).

177. Roux, J., Lefort, M., Bertin, M., *et al. Impact de la crise sanitaire de la COVID-19 sur la santé mentale des étudiants à Rennes, France* 2021. https://hal.science/hal-03172226 (2024).

178. Joanna. *Le SARS-CoV-2 : une expérience inédite de surveillance génomique mondiale* 2023. https://interstices.info/le-sars-cov-2-une-experience-inedite-de-surveillance-genomique-mondiale/.

179. *New York State Statewide COVID-19 Testing (Archived) | State of New York* https://health.data.ny.gov/Health/New-York-State-Statewide-COVID-19-Testing-Archived/xdss-u53e/about_data.

180. *Corona Zahlen: Karte für Deutschland + weltweit* https://interaktiv.morgenpost.de/corona-virus-karte-infektionen-deutschland-weltweit/.

181. *UKHSA data dashboard* https://ukhsa-dashboard.data.gov.uk/#category=nations&map=rate.

182. Dong, E., Du, H. & Gardner, L. An interactive web-based dashboard to track COVID-19 in real time. *The Lancet Infectious Diseases* **20,** 533–534. https://www.thelancet.com/journals/laninf/article/PIIS1473-3099(20)30120-1/fulltext (2020).

183. *About* https://www.global.health/about/.

184. *COVID-19 cases | WHO COVID-19 dashboard* https://data.who.int/dashboards/covid19/cases.

185. Khare, S., Gurry, C., Freitas, L., *et al.* GISAID's Role in Pandemic Response. *China CDC Weekly* **3,** 1049–1051. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8668406/ (2021).

186. Perra, N. Non-pharmaceutical interventions during the COVID-19 pandemic: A review. *Physics Reports* **913,** 1–52. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7881715/ (2021).